Luis Fernando Lara

Curso de lexicología

EL COLEGIO DE MÉXICO

CURSO DE LEXICOLOGÍA

CURSO DE LEXICOLOGÍA

Luis Fernando Lara

EL COLEGIO DE MÉXICO

413.07
L318c

Lara, Luis Fernando
Curso de lexicología / Luis Fernando Lara. -- México, D.F. : El Colegio de México, 2006.
248 p. ; 22 cm

ISBN 968-12-1187-1

1. Lexicología -- Estudio y enseñanza.

Primera edición, 2006

D.R. © El Colegio de México, A.C.
Camino al Ajusco 20
Pedregal de Santa Teresa
10740 México, D.F.
www.colmex.mx

ISBN 968-12-1187-1

Impreso en México

ÍNDICE GENERAL

Prólogo, 9

PRIMERA PARTE
LA DETERMINACIÓN DE LA UNIDAD *PALABRA*

1. La palabra fonológica, 17

2. La unidad de denominación, 37

3. La palabra morfológica, 53

4. El significado de la palabra, 85

5. La palabra escrita, 111

6. De la unidad palabra al vocablo, 133

SEGUNDA PARTE
EL LÉXICO Y SU ESTUDIO

Introducción: el léxico y sus abordajes, 143

7. La recolección de datos léxicos, 147

8. Estudios cuantitativos del léxico, 165

9. El estudio cualitativo del significado léxico, 181

10. El léxico, símbolo social, 213

11. La etimología, 231

Índice analítico, 249

PRÓLOGO

La lingüística es una ciencia dedicada a estudiar todas las lenguas del mundo; tanto en su sistematicidad y sus características internas, como en sus raíces sociales, históricas y culturales, y en su residencia en la mente humana. Puesto que las lenguas son fenómenos manifiestos para cualquier persona —los más exclusivamente humanos— y ofrecen una variedad asombrosa, la única manera de llegar a conocerlas es mediante su observación, la aplicación a sus manifestaciones de diferentes criterios y métodos de análisis, la producción de hipótesis que busquen comprenderlas y, finalmente, la emisión de teorías que las expliquen. La lingüística debe considerarse como una ciencia empírica, es decir, una ciencia basada en datos recabados de la experiencia. En cuanto ciencia, los resultados de sus investigaciones deben poderse verificar independientemente de quien los haya producido o de quien haya propuesto la manera de estudiarlos.

Desde hace más de medio siglo todas las corrientes de la lingüística, tanto las que se orientaron hacia la descripción de lenguas y sus instrumentos de estudio, como las que buscan proporcionarnos un modelo del funcionamiento de cada lengua o del funcionamiento de la facultad de hablar que caracteriza a los humanos, han tomado la *palabra* como elemento de los sistemas lingüísticos (central para la fonología, la morfología y la sintaxis), aunque poniendo en duda su existencia real como *unidad* verbal en favor, sobre todo, del morfema y el sintagma. Lo común en un curso de lingüística es usar palabras, usar la palabra *palabra* para darse a entender, pero desestimarla como fenómeno unitario y disolverla en la morfología. La unidad *palabra* viene a quedar reducida en los cursos a una extraña manía verbal de los hablantes de lenguas como el español o, en general, de las lenguas de Occidente en cuyas culturas se inventó la ciencia del lenguaje, que no tiene más definición que el aparecer entre dos espacios en blanco en un texto escrito —como

aquí— o entre dos pausas al hablar. (Si al hablar hiciéramos esas pausas, hablaríamos como telégrafo o como robot telefónico de las centrales administrativas.)

Este *Curso*, por el contrario, parte de la comprobación de que los hablantes de muchas lenguas tienen signos para designar las unidades *palabra*, como *word* en inglés, *Wort* en alemán, *mot* en francés y en catalán, *parola* en italiano, *vox* en latín, etc. y que esos signos se conformaron en la experiencia intelectual de los hablantes de estas lenguas, para distinguir esa clase de unidades lingüísticas discretas, claramente perceptibles para sus hablantes, antes de que se consolidara el pensamiento gramatical y se inventara la lingüística. Corresponde mejor a la seriedad de una ciencia tomar en cuenta todos sus datos, como el de la existencia de estas palabras para nombrar esta clase de signos, que atribuir un serio error conceptual a tantas generaciones de seres humanos que, por siglos, han hablado de ellos. Los signos *palabra* se nos presentan como primeros síntomas de la existencia real de las palabras, a cuyo análisis y teorización habremos de dedicarnos en la primera parte del *Curso*.

Debido a la manera en que, durante todo el siglo XX y lo que va del XXI, la lingüística se ha dedicado mayormente a investigar qué clase de sistemas son las lenguas, y ha descubierto que unidades como el morfema y las combinaciones de la sintaxis muestran con mayor facilidad sus características sistemáticas, el papel de la *lexicología*, o disciplina dedicada al estudio, precisamente, de las unidades *palabra* y de su conjunto, llamado *léxico*, ha quedado mal definido en el programa descriptivo de las lenguas y ha tendido a disolverse en la morfología, como *estudio de la formación sistemática* (no histórica, no genética, no etimológica) *de estos signos*; es decir, se ha considerado la lexicología como una *Wortbildungslehre* (como se llama en alemán —el término con mayor prosapia en la historia de la lingüística— a la *Formación de palabras*). El efecto de esta disolución de la palabra en la morfología ha sido la publicación de una docena o más de textos dedicados a la morfología o la formación de palabras, en los que la unidad *palabra* se da por sentada, sin merecer más de una o dos páginas de comentarios, y sin que haya habido una consideración dilatada de su probable existencia y del conjunto de sus características. Este *Curso*, por el contrario, se propone recuperar el lugar que le corresponde en el edificio de la ciencia del lenguaje.

En México, en donde la riqueza y la diversidad de las lenguas es tan asombrosa como su diversidad natural; en donde se habla una lengua europea, como lo es el español, que tiene esparcidos por Europa e Hispanoamérica varios núcleos de irradiación que le dan diversidad y personalidad distinta; y en donde las lenguas amerindias nos confrontan, en cuanto seres humanos y ciudadanos, con la tragedia de su riqueza cultural y su pobreza existencial, uno, como lingüista, siente la obligación moral y científica de contribuir a la formación de jóvenes colegas, que sean capaces de acercarse a todas sus lenguas, sin discriminación de unas en favor de otras, y observarlas, estudiarlas, describirlas y ofrecerlas al conocimiento de la sociedad y al conocimiento universal. Se convence uno de que sólo así, con buenos métodos de trabajo y concepciones adecuadas de lo que son las palabras y el léxico, es capaz de elaborar los sistemas de escritura (cuando es el caso), las gramáticas y los diccionarios de las lenguas amerindias que tanta falta hacen para la educación de sus pueblos y la defensa de su derecho humano a hablar sus propias lenguas.

Pero si el hecho de que el autor de este libro sea lingüista mexicano y enseñe en México es el motivo inmediato para escribirlo, eso mismo no tiene por qué convertirse en una limitación del horizonte de explicación y de utilización de la obra. Muy probablemente le resultará extraño al lector europeo o estadounidense, más que al hispanoamericano, que un libro como este se centre en México y no adopte sus propias coordenadas, que son las que suelen definir el espacio de significación del discurso científico. Hay una buena razón para ello: la excentricidad hispanoamericana en relación con España y en relación con el mundo científico, unida a la experiencia diaria del contacto con las lenguas amerindias, que se puede tener en casi toda Hispanoamérica, nos ofrecen, mejor que a los europeos y los estadounidenses, un horizonte de relatividad que nutre mejor la percepción y la comprensión de los fenómenos lingüísticos —incluso la comprensión de los fenómenos del español o de las otras lenguas europeas—; una condición absolutamente necesaria para el aprendizaje de la lingüística. Por eso se puede afirmar que la teoría, los criterios y los métodos expuestos aquí deben ser válidos y verificables en cualquier otra parte del mundo.

Este *Curso* tiene por objetivo formar lingüistas; darles una conformación conceptual sólida y bien fundada, que les sirva para emprender

el estudio del léxico de cualquier lengua, así como para poder leer y ponderar obras dedicadas al mismo tema, procedentes de la variada oferta de la investigación lingüística contemporánea.

No es fácil escribir un curso de esta clase. Por más que conozca uno a sus posibles lectores, siempre duda de su propia capacidad para hacerse entender y para alcanzar un nivel homogéneo de explicaciones, que sean a la vez sencillas y llenas de sentido. No es un curso "ecléctico", que resuma puntos de vista de otros autores o que presente diferentes corrientes de la lexicología, sin juzgarlas ni comprometerse con ellas o, lo que es peor, descartándose a sí mismo en adocenado respeto a los otros autores. Por lo contrario, es un curso con una orientación definida, dirigido por sus dos preguntas centrales: ¿qué es una palabra? y ¿qué es el léxico?, como preguntas que cualquier persona se hace cuando se propone adentrarse en esas consideraciones. No puede uno enseñar lo que no entiende; y no puede uno creer que ha aprendido algo, si no sabe poner en práctica ese aprendizaje. Como decía a menudo Jean Piaget, aprender es *aprender a hacer*. Este *Curso* busca *enseñar a hacer*, con el objetivo de facilitar a los estudiantes, hispanohablantes y hablantes de lenguas amerindias, la práctica de la lexicología.

Decir "práctica" no es decir que falte "teoría"; todo lo contrario: la trama del curso es una teoría; una teoría de la palabra y del léxico, junto con los criterios y los métodos mínimos necesarios para verificarla en cualquier lengua que se someta a estudio. Se ocupa de aquellos *elementos* de la lexicología, sin cuyo conocimiento y su profunda comprensión, la investigación contemporánea deja de tener sentido o conduce a la trivialización.

La primera parte del *Curso* trata sistemáticamente y en orden cada uno de los elementos y criterios que sirven a definir lo que es una palabra en cualquier lengua, por lo que supone poder ofrecer una teoría de la palabra, que después se sustancie en cada lengua particular. En su segunda parte se ocupa del léxico como conjunto de palabras, tratando de exponer sus relaciones con la memoria humana, sus características cuantitativas y su papel en la vida social, junto con los conceptos y los métodos de investigación de la semántica léxica.

En cada capítulo hay recuadros dedicados a subrayar los conceptos que es necesario comprender, y otros en los que se agrega información que les da contexto histórico y cultural. Al final de cada capítulo se ofrece una lista de los conceptos cuyo dominio se espera haber logrado después de su estudio. El libro no lleva notas de pie de página —salvo aquellas dedicadas a agradecer ciertos ejemplos a sus autores—, con objeto de no distraer su estudio con discusiones que más bien corresponden al tratado científico dirigido a colegas investigadores. Las obras que se listan al final de cada capítulo son sólo aquellas que han servido para diferentes desarrollos de ideas en el *Curso*. A lo largo del texto se mencionan varias obras más, cuya lectura es recomendable (incluso cuando lo que se sostiene en ellas es contrario a lo que propone el *Curso*), pero no se citan en la bibliografía.

El *Curso* supone la dirección de un profesor que ayude a comprenderlo y estudiarlo; supone también que el estudiante tiene los conocimientos gramaticales básicos, que se reciben en la escuela elemental, para comprender la terminología tradicional de la lingüística. En cambio, no supone haber cursado varias materias previas de lingüística, aunque conviene haber llevado previamente una de fonética y una de fonología.

Agradezco de todo corazón al Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra, de Barcelona, la beca que me concedió para pasar en su seno el año sabático 2003-2004; en particular a quien fue su directora, María Teresa Cabré, su generosidad y comprensión durante ese año. A ella también le agradezco su lectura y comentarios, así como a mis amigos Klaus Zimmermann, de Bremen, y Roberto Ham, de El Colegio de la Frontera Norte, la lectura, corrección y discusión de diversos tramos del libro. También les agradezco a mis compañeros de trabajo del *Diccionario del español de México* y a Martha Elena Venier, de mi Centro de Estudios Lingüísticos y Literarios, su paciencia para leer cada capítulo y aconsejarme en varios pasajes. Agradezco a Maira Colín su ayuda en la preparación del texto definitivo. Elizabeth Luna Traill, de la Universidad Nacional Autónoma de México, me hizo el honor de dictaminar favorablemente este trabajo.

Tepoztlán, 5 de noviembre de 2004

PRIMERA PARTE
LA DETERMINACIÓN DE LA UNIDAD *PALABRA*

"La palabra: este término vituperado... e irremplazable".

ÉMILE BENVENISTE

1
LA PALABRA FONOLÓGICA

INTRODUCCIÓN

Pocas unidades de la lengua parecen ser tan evidentes para cualquier persona que haya pasado por la escuela elemental como la *palabra*. No sucede lo mismo con la unidad *fonema* o con la unidad *morfema*. Por ejemplo, si intentamos que un amigo nos explique cómo es posible que un andaluz, un cubano o un jarocho "se coman las eses", es decir, las aspiren (particularmente antes de una consonante) y, sin embargo, seamos capaces de saber que se trata de eses, le crearemos un difícil problema conceptual, que muy probablemente no sea capaz de resolver, ya que hay que estudiar en la universidad y haber tomado un primer curso de fonética para saber en qué consiste la unidad que subyace a las eses que pronunciamos muchos hispanohablantes antes de una consonante (como en *pasto* o *casco*), y a las que realmente no pronuncian aquellos andaluces, cubanos o jarochos, sino que las aspiran o, en muchos, casos, las pierden por completo: [pahto, pą:to], [kahko, ką:ko] (y las *aes*, con una apertura): aprendemos que se trata de esa unidad abstracta (aunque seguramente tiene una naturaleza cognoscitiva hasta ahora poco estudiada) que llamamos *fonema*.

O consideremos el esfuerzo que tiene que hacer un maestro de escuela para que sus alumnos aprendan lo que es un *morfema*, y sepan distinguir *afijos de flexión* y *afijos derivativos*. En cambio, cuando hablamos de *palabras*, todos parecemos entender con facilidad de lo que estamos hablando, e incluso podemos aducir muchísimos ejemplos de ellas: *pasto*, *casco*, *perro*, *correr*, *mirábamos*, *contra*, *según*, *pero*, *la*, *con*, etcétera.

Y, sin embargo, para una buena cantidad de lingüistas la unidad *palabra* es un fenómeno de tan difícil delimitación, que incluso prefieren declararla "precientífica", "objetivamente inexistente", o por lo menos inútil. Por ejemplo, *La lingüística. Guía alfabética*, obra de referencia di-

rigida por el gran lingüista francés André Martinet, en su artículo 38, después de exponer las dificultades y las propuestas que se han hecho para definirla, se pregunta "si, en el campo de la investigación, no sería más interesante eliminar completamente el término palabra y sustituirlo por monema y sintagma, según los casos, que tienen la ventaja de ser utilizables con referencia a todas las estructuras lingüísticas", aunque agrega en seguida: "No obstante, es innegable que en el plano psicológico, la palabra es un elemento real y válido; el sujeto hablante medio, aun inculto, y, según Sapir, cualquiera que sea la estructura de su lengua, piensa y pronuncia 'palabras', de manera que los propios lingüistas se ven obligados a utilizar en sus escritos la palabra en el sentido tradicional y habitual".

¿A qué se debe que la unidad *palabra* parezca ser tan evidente para muchísimas personas, tan "válida" psicológicamente hablando y, sin embargo, de tan difícil explicación lingüística? En los siguientes cinco capítulos nos ocuparemos de analizar este problema y de buscarle una solución que, a pesar de serlo, no resultará sencilla.

Influencia de la tradición escrita

Lo primero que tenemos que despejar es el papel que tiene la idea "precientífica" o "psicológica" de la palabra en nuestra idea de las lenguas. Para hacerlo hay que tomar en cuenta dos situaciones históricas importantes de la comunidad lingüística hispánica: en primer lugar, que el español tiene una historia de reflexión acerca de sí mismo, como lengua, ya cercana al milenio completo, gracias a la cual hemos aprendido a objetivarlo y analizarlo de una manera tan detallada y tan profunda, que nos convierte en parte de un pequeño sector de la humanidad que dispone de una lengua muy cultivada y muy claramente objetivada, como se puede observar en el gran número de gramáticas, diccionarios, ortografías, historias, etc. con que contamos; en segundo, que el español se escribe desde la Baja Edad Media y ha venido refinando una idea de palabra gráfica en la que hemos sido educados, mediante la cual la lectura se ha vuelto un procedimiento muy eficaz de comunicación verbal.

Aquello que nos permite afirmar que sabemos lo que es una palabra es producto de esos dos fenómenos culturales de la lengua española. Los

diccionarios, particularmente, nos han acostumbrado a considerar que los conjuntos de trazos gráficos continuos (es decir, sin separaciones notables entre las letras), que encabezan los artículos lexicográficos, generalmente con un tipo de letra diferente, negrita o de mayor tamaño que el resto, y que se listan en orden alfabético, *son palabras*.

La escritura, por su parte, en cuyo trazo y manejo se nos educa en la escuela, nos ha acostumbrado también a pensar que las palabras son configuraciones gráficas separadas por espacios en blanco, razón por la cual son tan palabras *perro*, *atmosférico* y *quisiéramos*, como *el*, *de* e *y*.

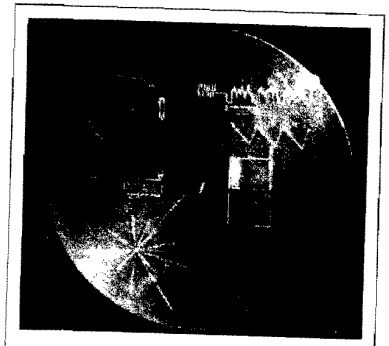
Sin duda el concepto gráfico de palabra, que recibimos de nuestra tradición reflexiva lingüística, tiene un papel determinante en nuestra idea de la *palabra*, que habrá que tomar en cuenta, pero si la lingüística es una ciencia que se ocupa de conocer todas las lenguas, y la gran mayoría de ellas no cuenta, hasta la fecha, con una cultura de reflexión y una escritura como la nuestra, es necesario buscar en qué consiste esa supuesta existencia universal de la palabra, como afirma Martinet que lo cree Sapir, sin referirla a la palabra gráfica del sistema alfabético de escritura.

Es necesario actuar así, además, porque es la mejor manera de entender las características de la lengua propia que, como lengua, no difiere de las del resto de la humanidad. Es decir, tenemos dos buenas razones para no contentarnos con el concepto gráfico de la palabra: una de disposición intelectual para abrirnos a los muy variados fenómenos lingüísticos del género humano, que son mayoritariamente hablados y no escritos, y otra de capacidad metódica para poder delimitar una unidad que parece tener tanto valor organizativo y cognoscitivo en todas las lenguas.

1. CONCEPTOS BÁSICOS

Todos los métodos de descripción de lenguas operan mediante análisis e imponiendo a cada lengua considerada una *organización por niveles*, que parecen corresponder a las características internas de todos los sistemas lingüísticos. Para comprender el porqué de estos procedimientos, imaginemos lo que sucedería en el momento en que realmente aterrizara en nuestro planeta un ser de otro mundo.

Lo primero que nos interesaría descubrir es si tiene algún medio que le permita comunicarse con nosotros, pues de esa manera podríamos ofrecerle una recepción pacífica y entablar con él un diálogo enormemente fructífero, ya que no sólo le preguntaríamos de dónde viene y por qué viene, sino además cómo logró llegar desde tan lejos; nos enteraríamos de las características de su especie y de la civilización de la que procede; le contaríamos lo que somos o queremos ser los humanos; lo introduciríamos en el aprecio de nuestra naturaleza y de nuestras culturas, etc. (una recepción contraria, a balazos, poseídos de un temor histérico, es lo que acostumbramos ver en el cine hollywoodense, pero esperemos que nunca llegue a hacerse realidad).



Los científicos que lanzaron al espacio las sondas llamadas Voyager (I y II), en la década de 1970, se plantearon cómo enviar un mensaje terrícola a algún ser inteligente que se topara con esas sondas en la profundidad de nuestra galaxia. He aquí la reproducción de la portada del disco que contiene instrucciones para leerlo, en cuyo interior hay películas de la Tierra y grabaciones de muchas lenguas, de música, etc. Las instrucciones básicas, en código binario, sirven para determinar la velocidad con que se debe tocar el disco, sobre la base de un período constante, concebido como universal, correspondiente a la transición de un estado a otro de un átomo de hidrógeno. La "estrella" que se ve abajo a la izquierda es un diagrama de la localización de la Tierra, determinado por el cruce de 14 quasares.

Para poder descubrir si ese ser de otro mundo tiene algún medio de comunicación que podamos utilizar para conversar con él, tendremos que comenzar por buscar acontecimientos perceptibles para nuestros órganos, emitidos por ese extraterrestre, que sigan algunos patrones repetidos, a base de los cuales podamos hacer un inventario posible de los elementos que forman su sistema de comunicación. Tal tarea, según nos cuenta la prensa, se lleva a cabo desde hace ya muchos años en algunos observatorios astronómicos, equipados con radiotelescopios, como el de Arecibo, en Puerto Rico, en los que se buscan patrones sistemáticos de emisiones de algún tipo de radiación, pongamos por caso, de radio, para proceder a clasificarlos, a reconocer la manera en que se combinan y, posteriormente, a buscarles algún sentido.

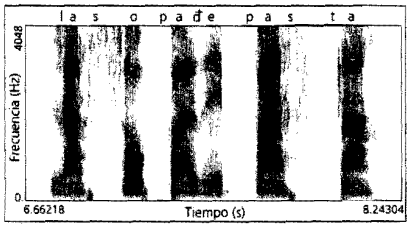
El astrónomo Carl Sagan, incluso, logró enviar un mensaje "terrícola" de esta clase al espacio ignoto en las sondas espaciales Voyager, en la década de 1970, y escribió, con el mismo tema, una interesante novela, después llevada al cine, con el título de *Contacto*.

Supongamos que ese medio de comunicación es sonoro, como nuestras lenguas. Una comunicación de esa clase, que ofrezca patrones repetidos de estímulos, se desarrolla en el tiempo, como una sucesión. Tomaremos, por lo tanto, trechos tan grandes como sea posible de sucesiones sonoras, que procesaremos en un laboratorio de acústica, para poder estudiar las características de ese sonido y tratar de segmentarlo en elementos que, por repetirse, consideramos compositivos de su lengua.

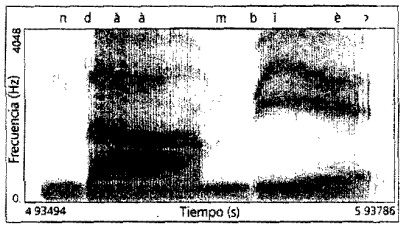
No es nada fácil hacer una segmentación de esa clase porque, de entrada, no sabemos en dónde se puede segmentar el continuo sonoro,[1] de manera que lo que haríamos sería efectuar algunas segmentaciones arbitrarias, que contuvieran segmentos sonoros iguales repartidos en ellas, y que empezaríamos a comunicar a nuestro visitante extraterrestre para observar sus reacciones. Este sería un *primer nivel de análisis*, que podríamos llamar de *inspección fonética*.

En realidad, un procedimiento de esta clase es el que sigue cualquier lingüista que se enfrenta, por primera vez, a una lengua humana que le es desconocida. Con la diferencia, muy importante, de que sabemos que todo ser humano habla (salvo trágicos casos de pérdida del habla o de mudez), y que el habla consiste de una cadena de unidades discretas, que se combinan sistemáticamente en la sucesión sonora: los fonemas.

Así que, en el *nivel de inspección fonética, primer nivel de análisis lingüístico*, en donde tomamos en cuenta las peculiaridades acústicas y per-



Espectrograma de "la sopa de pasta".



Espectrograma de "un borrego" en mixteco.

[1] Como se ve en estos dos espectrogramas, que agradezco a la doctora Esther Herrera, del Centro de Estudios Lingüísticos y Literarios de El Colegio de México.

ceptuales del sonido de las lenguas, podemos hacer una detallada clasificación de los sonidos y sus combinaciones en la lengua estudiada, aun sin saber, todavía, ni cuáles son, como unidades, ni lo que ayudan para averiguar *lo que se quiere decir*.

Es posible hacer listas de las combinaciones de los sonidos encontrados de dos en dos, de tres en tres, etc. Podemos también suponer qué sonidos serán consonánticos y cuáles vocálicos, puesto que la diferencia entre ambas clases es de carácter acústico y articulatorio, pero no ir más allá, pues de ese análisis es imposible discernir qué sonidos son manifestaciones de qué fonemas; es decir, no podemos reconocer una *identidad* de los sonidos; algo que nos permita, por ejemplo, reconocer que tanto una [s] como una [h] o la pronunciación alargada y abierta de una vocal en varios dialectos andaluces, cubanos o el jarocho, son manifestaciones del fonema /s/.

1.1. La búsqueda de las invariantes

Es un problema de las ciencias en general poder encontrar invariantes en una naturaleza intrínsecamente variada. Por ejemplo, consideremos los frijoles: hay una gran cantidad de semillas leguminosas, que crecen en vainas y que tienen variadas formas ovaladas o de riñón. Varían en tamaño, color, región, clima, época, etc. según donde crecen; unos son comestibles y otros hacen daño. ¿Qué es lo que hace que los campesinos distingan los frijoles alimenticios, que nos interesan, de muchas otras plantas que "parecen" frijoles o que, incluso botánicamente, "sean" variedades de frijol (como las semillas del colorín o las del flamboyán, por ejemplo)? El problema es de *identificación de una invariante*, el frijol comestible, entre muchas variedades.

Para explicar el mismo tipo de problema en lingüística, Ferdinand de Saussure, padre de la lingüística moderna, propuso su famosa distinción entre *lengua* y *habla*. Para Saussure, el *habla*, que es la realidad infinitamente variada del hablar una lengua, corresponde a la variación, en tanto que la *lengua* es la invariante que buscamos identificar en el habla. Poder llegar a identificar las invariantes que subyacen al habla y constituyen el sistema de una lengua es una *condición de posibilidad de la lingüística*; es decir, no se puede hacer el estudio lingüístico de un idioma si no se logran identificar las invariantes subyacentes en el habla.

Para averiguar esto, es necesario pasar a un segundo nivel de análisis, que nos conduzca a la *identificación de los fonemas*, es decir, al paso a las *invariantes*, al paso a la *lengua*, pero, desgraciadamente para nosotros, a quienes nos gustaría que el análisis fuera de la misma clase que antes y se pudiera llevar a cabo con verdadero automatismo (más ahora, que contamos con el maravilloso instrumento que es la computadora electrónica), el segundo nivel de análisis requiere tomar en cuenta el significado de los signos o, dicho más llanamente, requiere la intervención del significado.

1.2. La teoría del signo

La segunda de las enseñanzas de Saussure es su *teoría del signo lingüístico*. Para él, no hay signo lingüístico si no está compuesto, de manera inseparable, de un *significante* y un *significado*. El *significante*, o *plano de la expresión* del signo (según el modo en que el gran lingüista danés Louis Hjelmslev explicaba a Saussure), consiste, ante todo, en el sonido o en la materia gráfica, cuando se trata de escritura, del signo: aquello que hemos logrado clasificar en nuestro primer nivel de análisis. Lo llamaremos, con Hjelmslev, *sustancia de la expresión*. A esta sustancia, infinitamente variada en su realidad acústica y articulatoria (y gráfica), corresponde una *forma de la expresión*, constituida por las invariantes que subyacen a la sustancia, los fonemas (y las letras del alfabeto). Pero para poder reconocer al significante necesitamos el significado, pues de otra manera, la sustancia de la expresión no es otra cosa que sonido ininteligible. Tomemos un ejemplo de la historia de la poesía surrealista:

"Sonido poético":[2]
gadji bim bimba
glandridi lauli lonni cadori
gadjama bim beri glassala...

[2] Tomado de Raúl Gustavo Aguirre, *El dadaísmo*, Buenos Aires, Centro Editor de América Latina, 1968, p. 48.

Su autor, el poeta y pintor surrealista-dadaísta Hugo Ball, creía poder alcanzar así la pureza poética del sonido, en la creencia de que lo esencial del hablar es el sonido de la voz humana. Lo que obtuvo fue la imposibilidad total de comprender esa poesía; es decir, la destrucción del signo lingüístico. Esas combinaciones de letras, que nos ofrece como si fueran palabras, no son nada, en ninguna lengua. Son *materia* sonora o gráfica que parece lengua, que incluso "suena", por haber utilizado letras del alfabeto latino y construcciones gráficas como las de las palabras, pero no son sustancia de la expresión de ningún signo; no son signos; no son verdaderas *palabras*: no tienen significado.

Una buena cantidad de métodos de descripción de lenguas, en particular los que provienen de las enseñanzas del lingüista estadounidense Leonard Bloomfield, hacen intervenir el significado en el segundo nivel de análisis, en el nivel fonológico, para poder reconocer la identidad de los fonemas, aunque no reconocen la inseparabilidad de *significante* y *significado* como lo hace Saussure; en especial, para Bloomfield, el significado solamente tiene valor de apoyo, de "muleta" para la descripción o, dicho con una palabra más científica: *heurístico*.

Para la lingüística que proviene de la enseñanza de Saussure, en cambio, el significado es una parte plenamente constitutiva del signo, como se verá más adelante en este *Curso*, cuando se desarrolle el *plano del contenido* del signo lingüístico.

1.3. El método de la conmutación

La teoría saussureana del signo da lugar a un *método de descubrimiento* muy importante, no solamente para este nivel de análisis, sino para el resto del análisis lingüístico: el *método de la conmutación*. Éste consiste en oponer al menos dos *signos* entre sí, de cuyo *significante* sólo hacemos variar un elemento sonoro, que hipotéticamente consideremos manifestación de un fonema. Por ejemplo:

[pérro]: [péso]

Como se ve, lo único que difiere entre estas dos emisiones sonoras son dos hipotéticos fonemas (hipotéticos, porque estamos haciendo

como si no supiéramos español): /rr/ vs. /s/. Si la comparación entre los dos significantes da lugar a una diferencia de significado, la diferencia entre /rr/ y /s/ será *pertinente* y demostrará que se trata de dos fonemas distintos. En cambio, si oponemos

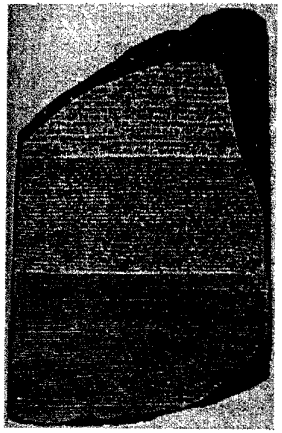
[káhko]: [kásko]

La hipótesis de que haya, en español, un fonema /h/ se falsifica, pues ambas pronunciaciones no dan lugar a la distinción de dos significados. Se trata, en los dos casos, de dos *sonidos* distintos —resultado del primer nivel de análisis—, correspondientes a un mismo fonema /s/, que viene siendo un resultado del segundo nivel de análisis.

Podemos esquematizar el procedimiento de la conmutación:

significado $_a$	3er. paso ↔	significado $_b$
2o. paso ↕	**comparación**	↕
significante $_a$	1er. paso ↔	significante $_b$

El orden de estos pasos es resultado del método de descubrimiento que tenemos que seguir al enfrentarnos a una lengua desconocida: primero comparamos dos posibles significantes a y b, uno de cuyos fonemas hipotéticos varía ([pérro]: [péso], [káhko]: [kásko]); después consideramos sus correspondientes significados a y b ('perro': 'peso', 'casco'), y, si encontramos que hay diferencia de significado (tercer paso), entonces tenemos signos distintos, a partir de los cuales discernimos la *identidad* de uno de los fonemas. Después de proceder a descubrir todos los fonemas de la lengua mediante sucesivas conmutaciones (método conocido como de *pares mínimos*), habremos identificado todos los fonemas y habremos descrito la *forma de la expresión del signo*; es decir, habremos descubierto las invariantes que son pertinentes en el sistema fonológico de la lengua en cuestión (y las variedades sonoras correspondientes a los mismos fonemas ayudarán a definir distintos dialectos del mismo sistema). Lo que al principio, en el primer nivel de análisis se nos presentaba como *materia sonora*, se ha vuelto *sustancia de*



Nos sirve para ejemplificar este proceso el caso histórico de la PIEDRA DE ROSETTA, fragmento de una estela de 95 × 76 cm, encontrada por el ejército de Napoleón cuando invadió Egipto en el siglo XVIII, en la ciudad de Rashid o Rosetta, y llevada después como botín de guerra al Museo Británico (cuando los ingleses derrotaron a Napoleón en Egipto). Se trata de un documento histórico del año 302 a.C., escrito en tres sistemas de escritura diferentes: en su parte alta, con jeroglíficos egipcios; en medio, en escritura demótica, una forma desarrollada a partir de los jeroglíficos egipcios manuscritos llamados "hieráticos", y abajo, en griego antiguo, en escritura griega. El joven francés de 18 años Jean-François Champollion logró descifrar el texto egipcio gracias al fragmento en griego antiguo, una lengua que conocía. Es decir, el griego antiguo le dio el sentido, la clave, el significado para poder descifrar la escritura demótica y, con ella, llegar a conocer los textos jeroglíficos del antiguo Egipto.

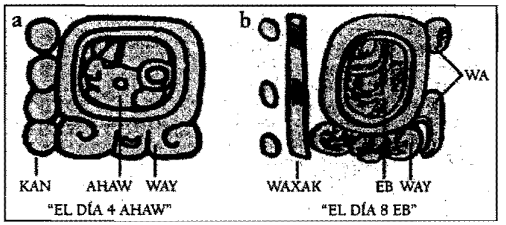
Tomado de la Enciclopedia libre *Wikipedia*

la expresión. Los resultados obtenidos constituyen el *segundo nivel de análisis*, que llamaremos *de análisis fonológico*.

Tener que apelar al significado para poder identificar los fonemas y, después, identificar los signos de una lengua implica, bien un hablante de la lengua que nos vaya diciendo, para cada par mínimo, si hay o no cambio de significado, o bien, al menos, alguien o algún registro que nos ofrezcan una clave de interpretación de los signos de la lengua, que nos permitan ir deduciendo el resto.

En cambio, pensemos en las dificultades todavía no bien resueltas que tienen los arqueólogos mayistas con las inscripciones de las ruinas mayas. Para la mala fortuna de los interesados en la escritura maya, no hay una "piedra de Rosetta" escrita en jeroglíficos mayas ni algún otro sistema de escritura que conozcamos, que les pueda dar la clave de interpretación del significado de esos signos. Han tenido que sumar muchos esfuerzos para llegar a las interpretaciones actuales, que les permiten ofrecer hipótesis probables acerca de lo que dicen los monumentos y las estelas del mundo maya. Hoy en día, gracias a una identificación de los jeroglíficos correspondientes a los meses del año, que logró fray Diego de Landa en el siglo XVI (afortunadamente antes de que la Conquista destruyera para siempre la continuidad de la memoria histórica de los pueblos mayas), gracias también a que el sistema de representación numérica es de más fácil reconocimiento y, por último, gracias a la hipótesis de que hay una relación

silábico-fonológica entre los jeroglíficos y varias lenguas mayas cuyo pasado se ha reconstruido (claro, hipotéticamente), como el chol en el caso de las inscripciones de Palenque, se piensa que la comprensión del significado de esa escritura va aumentando y mejorando, al punto de que se han podido establecer, por ejemplo, genealogías de los reyes de Palenque y otras regiones mayas.



En estos dos glifos de una estela de Palenque se puede reconocer la representación de los números a la izquierda. *Way* significa 'día de 24 horas'; *ahaw* y *eb* son nombres de días del *ts'olk'in*, el ciclo ritual adivinatorio de 260 días.

Los glifos provienen del artículo "Glifos y representaciones mayas del mundo subterráneo", de Guillermo Bernal Romero, *Arqueología Mexicana*, 8, 48 (2001), pp. 42-47.

2. SEGUNDO NIVEL DE ANÁLISIS: ANÁLISIS FONOLÓGICO

El caso es que nuestro segundo nivel de análisis nos ofrece una identificación de fonemas, que podemos organizar en un *cuadro fonológico*:

En el caso del español mexicano:

Consonantes					
	labiales	*dentales*	*alveolares*	*palatales*	*velares*
oclusivas sordas	p	t		ch	k
sonoras	b	d		y	g
fricativas sordas	f		s		x
nasales	m		n	ñ	
lateral			l		
vibrantes			rr/r		

Vocales			
	anteriores	*media*	*posteriores*
cerradas	i		u
abiertas	e		o
central		a	

En el caso del tzeltal de Oxchuc, Chiapas:

Consonantes						
	bilabiales	*alveolares*	*alveopalatales*	*palatales*	*velares*	*glotales*
oclusivas sordas	p	t			k	ʔ
oclusivas sordas glotalizadas	p'	t'			k'	
oclusiva sonora	b					
fricativas sordas		s	š			h
africadas sordas			ts	č		
africadas sordas glotalizadas			ts'	č'		
nasales	m	n				
laterales		l				
semiconsonantes	w			y		

Vocales			
	anteriores	*central*	*posteriores*
altas	i		u
medias	e		o
baja		a	

Los cuadros fonológicos, que identifican los fonemas de una lengua y establecen las relaciones estructurales entre ellos, son una *primera condición* para poder proceder a determinar la palabra en cada lengua.

2.1. La palabra como simple combinación de fonemas

Los fonemas de una lengua se combinan para ofrecer al signo su *forma de la expresión*. Pero no se combinan todos entre sí, como podría uno creerlo. Si, por ejemplo, los 22 fonemas del español mexicano se combinaran sin ninguna restricción, el número de combinaciones posibles, representado con la expresión matemática *n*! (*ene* factorial), sería el resultado de la multiplicación de $22 \times 21 \times 20 \times 19 \times \ldots \times 2 \times 1$, lo cual daría una cifra astronómica de posibles cadenas de fonemas y, en consecuencia, de palabras posibles.

Pero imaginemos la imposibilidad que tendría un ser humano para guardar en su memoria todas las palabras que se formaran de esa manera en su lengua y, en consecuencia, lo poco económica que resultaría la lengua. De ahí que todas las lenguas del mundo pongan ciertas restricciones a las cadenas de fonemas que podrían formar, restricciones que actúan precisamente en la determinación de la forma de sus palabras.

Esas restricciones se pueden expresar con la fórmula matemática de las permutaciones: si se combina el número de fonemas de la lengua (n), con el tamaño de la cadena (r), tenemos:

$$P(n, r) = \frac{n!}{(n-r)!}$$

De tal manera que las permutaciones posibles de los fonemas del español en cadenas de dos elementos serán 22!/ (22-2)! = 22! / 20! = 462 (y darían formaciones como *ba*, *en*, *fi*, *pl*, pero también *tk*, *fs*, etc.). Para formar una cadena de tres elementos serán 22!/ (22-3)! = 9 240, y así sucesivamente.

Pero aun cuando de esta manera se reduciría bastante la primera cifra de combinaciones posibles de los fonemas, el número de palabras del español, concebidas sólo como cadenas de fonemas, sería extremadamente alto.

2.2. La sílaba

Otra restricción a la formación fonológica de las palabras tiene su origen en las características del aparato fonatorio humano: puesto que las consonantes son modificaciones o interrupciones de la corriente de aire procedente de nuestros pulmones, que se producen en el interior de la boca, para que puedan "sonar" muchas de ellas necesitan llevar una vocal antes o después de su pronunciación. Por ejemplo, para poder pronunciar una consonante oclusiva, es decir, en la que se impide por completo la salida del aire, como /p/ o /t/, hace falta una vocal después de ellas: [pe], [te]. Tal restricción da origen a la *sílaba*, una unidad natural de todas las lenguas, en que su *núcleo* es una vocal y su *margen* es una consonante (por eso las vocales, que se pueden pronunciar solas, suelen valer también como sílabas).

Cada lengua dispone de un conjunto limitado de formaciones silábicas, debido a las características de sus fonemas. Tal conjunto de formaciones o estructuras silábicas es lo que llamamos *patrón canónico de la sílaba* de la lengua en cuestión.

En el siguiente cuadro están las sílabas más frecuentes del patrón canónico del español, establecido de manera cuantitativa:

estructura silábica	*ejemplos*
C'V	me.cá.te, té.tri.co, co.lé.gio, etc.
CV	ca.bé.za, te.lé.fo.no, etc.
C'VC	mís.ti.co, co.me.rán, tál.co, cóm.pu.to, etc.
V'C	ác.to, ín.te.gro, áp.to, etc.
V'	á.la, é.se, í.pi.co, etc.
CVC	pa.rá.dos, dan.zár, cal.ce.tín, etc.
V	a.lién.to, e.mí.lio, u.sár

C representa cualquier consonante; V, cualquier vocal. En los ejemplos, la separación silábica se indica con puntos. La acentuación es prosódica y no ortográfica; tampoco la escritura; véase el capítulo 5, § 1.2.3.1. La estructura silábica ejemplificada está subrayada.

De las sílabas del tzeltal de Oxchuc, Chiapas, su patrón canónico (en orden de frecuencia) es:

estructura silábica	*ejemplos*	*glosa*
(')V	(')u.o.bál	'luna'
VVC	aan.él	'calumnia'
VC	al.jóm	'parturienta'
CV	ba.cho'ó	'tuza'
CVC	chan.póm	'culebra'
VCC	ánts	'mujer'
CVCC	pájch'	'piña'

2.3. Fonemas con función demarcativa

Una restricción más a la formación fonológica de las palabras es el hecho de que en una lengua ciertos fonemas sólo puedan ocupar determinadas posiciones en la sílaba o en una sucesión de sílabas. Por ejemplo, en japonés, en posición inicial de palabra no es posible ninguna combinación de consonantes; en posición interna de la palabra, sólo se admite la combinación de *n* más consonante; en posición final, sólo se encuentra vocal o vocal más *n*. Es decir, en el patrón silábico del japonés no es posible una formación CCV (a menos que la primera consonante sea *n*) y, por eso, una palabra japonesa no puede comenzar con dos o más consonantes, de modo que, cuando se llegan a encontrar dos consonantes juntas —una de las cuales no sea *n*—, entre las dos hay una frontera silábica, que a su vez se vuelve *frontera de palabra* (cuando un japonés comienza a aprender español, tiende a pronunciar la palabra *grande* como *gurande*, que ajusta a su propio patrón silábico: CV.CVn.CV).

De esas características de las estructuras silábicas de las lenguas deriva, en muchas de ellas, un carácter más de sus fonemas: el que puedan servir como marcas de comienzo o final de palabras, es decir, como fronteras de palabras; esa característica nos ofrece la *función demarcativa de los fonemas* de una lengua, tal como la define el filólogo español Emilio Alarcos Llorach.

La función demarcativa de los fonemas puede clasificarse mediante tres criterios:

a] Según su relación con la función distintiva del fonema puede ser fonológica o afonológica.

b] Según su carácter simple o complejo; es decir, si se logra mediante un solo fonema o con varios fonemas agrupados.

c] Según que los fonemas señalen con claridad los límites de la palabra (función demarcativa positiva) o sólo indiquen en dónde no puede haber límite (función demarcativa negativa).

Veamos ahora criterio por criterio: una oposición fonológica es *demarcativa fonológica* si la oposición se produce siempre en posición inicial o final absoluta de la palabra. Por ejemplo, según afirma Alarcos Llorach, en griego sólo puede aparecer espíritu áspero (una especie de aspiración) en posición inicial de palabra; en todas las demás posiciones, sólo se da el espíritu suave (no aspiración). En español hay una oposición fonológica que tiene función demarcativa, pero parcial: la que se da entre /rr/ y /r/, pues en posición inicial de palabra sólo puede aparecer el fonema /rr/ y no /r/ : /rróto/, /rrápido/; no es completamente demarcativa porque entre vocales aparecen /rr/ y /r/ : /yerro/ (hierro), /yero/ (hiero), así como aparece /rr/ antes de /n/, sin que se trate de una demarcación entre palabras: /en.rrí.ke/ (Enrique).

Un fonema tiene función demarcativa *afonológica*, en cambio, cuando es sólo un alófono el que determina la posición inicial o final de una palabra. Por ejemplo, en japonés, el fonema /g/ se realiza como [g] sólo en principio de palabra; si no, se realiza como nasal velar [ŋ]. Algo semejante se podría decir de las consonantes oclusivas sonoras en español (/b,d,g/), que sólo se pronuncian como tales en posición inicial absoluta de palabra, es decir, después de una pausa clara en la expresión (aunque también son oclusivas después de /m/, /n/ u otra oclusiva), pues en las otras posiciones se pronuncian fricativas: [#bjen], [#dáto], [#gáto], pero [está b̞jén], [sol̪d̪áð̪o], [lagárto].

Cualquier fonema aislado, que tenga función demarcativa, es *simple*; en cambio, una combinación de fonemas con función demarcativa es *compleja* cuando la primera parte de la combinación se da al final de la palabra, y la segunda al principio, pero nunca en su interior. Por ejemplo, en francés, una vocal nasal seguida de /m/ sólo se da entre palabras: *on mange*.

Una combinación es *demarcativa, compleja y afonológica* cuando se produce mediante la presencia de alófonos. Por ejemplo en alemán, el

fonema fricativo velar /χ/ se realiza como [χ] tras vocales posteriores: *kuchen*, *machen*, *kochen* (muy parecido al fonema del español /x/) y palatalizado, es decir, adelantado, tras las demás vocales y las labializadas: *küche*, *riechen* (como en la pronunciación chilena de *mujer*, *escoger*). Pero cuando se trata del morfema de diminutivo —*chen*, siempre se realiza palatalizado, independientemente de las vocales que lo antecedan: *hundchen*, *liebchen*; es decir, demarca al morfema.

Todos estos ejemplos son de señales demarcativas positivas; son negativos los fonemas que nunca aparecen en posición inicial o final de la unidad, como /r/ y los grupos /pt/ y /ks/ del español, o la [ŋ] del japonés.

2.4. Fenómenos prosódicos con función demarcativa

Todas las lenguas habladas consisten en un flujo de aire que sale de la boca, no solamente con diferentes clases de interrupciones, sino además con cierta "melodía" y con variadas combinaciones de mayor o menor intensidad en la emisión de la voz. En muchas lenguas del mundo se percibe un *acento de intensidad*, generalmente sobre diferentes sílabas de la expresión verbal. Este acento de intensidad puede ser pertinente para la significación, como sucede en español, en donde cada cambio del acento de intensidad hace que los signos tengan diferentes significados. Por ejemplo en:

depósito / deposito / depositó

Cuando la palabra es esdrújula, queremos decir 'lugar o edificio en donde se guardan objetos durante cierto tiempo, para protegerlos o para conservarlos'; cuando es grave, estamos manifestando la primera persona del singular del presente de indicativo del verbo *depositar*; cuando es aguda, en cambio, significamos la tercera persona del singular del pretérito del mismo verbo.

Hay lenguas en las que hay acento de intensidad, pero no es pertinente, es decir, no sirve para manifestar cambios de significado. En francés, por ejemplo, el acento de intensidad cae siempre en la última sílaba de la palabra:

"Je veux, pour composer chastement mes églogues
Coucher auprès du ciel, comme les astrologues".

CHARLES BAUDELAIRE
"Tableaux parisiens", en *Les fleurs du mal*

En esos dos versos, *veux* 'quiero' y *ciel* 'cielo' son monosilábicas; *composer* 'componer', *chastement* 'castamente', *églogues* 'églogas', *coucher* 'dormir', *auprès* 'a la vera' y *astrologues* 'astrólogos' llevan acento de intensidad en su última sílaba (los demás signos no llevan acento de intensidad, lo cual constituye un problema para saber si se trata de palabras o no, de lo que se hablará más adelante). El acento de intensidad, que siempre aparece en la última sílaba, es un indicador afonológico simple y positivo de *frontera final de palabra.*

Lo mismo sucede, por ejemplo, con el tzeltal de Chiapas. En *chojak* 'red', *chomba'jel* 'comerciante' o *alal* 'niño', el acento de intensidad no es pertinente, pero como cae en la última sílaba es un demarcador de frontera final de palabra. En checo, en cambio, el acento cae siempre en la sílaba inicial, por lo que puede indicar frontera inicial de palabra. E incluso en latín, lengua en la que hay acento de intensidad, pero no es pertinente, el acento cae en la penúltima sílaba, si ésta es larga: *dominúsque* 'y el señor' o en la antepenúltima, si la penúltima es breve: *dóminus* 'señor'.

Aun más, se puede dar como regla general que siempre que haya acentos de intensidad en una lengua, sean o no fonológicamente pertinentes, indicarán la existencia de palabras, diferenciadas entre sí por esos acentos.

3. LA PALABRA FONOLÓGICA

De este capítulo se puede llegar a la siguiente conclusión:

Desde el punto de vista fonológico —es decir, de la *forma de la expresión del signo*—, una *palabra* es una sucesión de fonemas de la lengua en cuestión, organizados en sílabas, según las restricciones fonéticas (articulatorias y acústicas) que determina su estructura fonológica. En diversas lenguas esa misma estructura fonológica da lugar a la existencia de una *función demarcativa* de ciertos fonemas, que ayudan a delimitar las fronteras entre palabras; igualmente sus elementos prosódicos, como el acento de intensidad, ayudan a determinar la forma de la palabra.

Sin embargo, estas que son *condiciones necesarias de la determinación de la unidad palabra, no son condiciones suficientes* para lograrlo.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

nivel de análisis; lengua y habla; variante e invariante; signo lingüístico; contenido y expresión; significante y significado; forma y sustancia; fonética; fonología; fonema; alófono; conmutación; par mínimo; cuadro fonológico; combinación; permutación; factorial; sílaba; elementos segmentales y suprasegmentales o prosódicos; función demarcativa de un fonema.

OBRAS CONSULTADAS

ALARCOS LLORACH, Emilio, *Fonología española*, Madrid, Gredos, 4a. ed., 1965.
HJELMSLEV, Louis, *Prolegómenos a una teoría del lenguaje* [1943], traducción de José Luis Díaz de Liaño, Madrid, Gredos, 1974.
KRÁMSKÝ, Jiří, *The Word as a Linguistic Unit*, La Haya, Mouton, 1969.
LYONS, John, *Introducción en la lingüística teórica*, traducción de Ramón Cerdá, Barcelona, Teide, 1977.
MARTINET, André (dir.), *La lingüística. Guía alfabética* [1969], Barcelona, Anagrama, 1975.
SÁNCHEZ GÓMEZ, Francisco J., "Proyecto para la elaboración del diccionario monolingüe tzeltal: justificación teórica y metodológica", tesis, México, CIESAS, 2003.
ZAPATA GUZMÁN, Alfredo, *et al.*, *Swejteseji'bal k'op yu'un tseltal*, Tuxtla Gutiérrez, Gobierno del Estado de Chiapas, 2001.

2
LA UNIDAD DE DENOMINACIÓN

INTRODUCCIÓN

En el capítulo anterior definimos el carácter fonológico de la unidad *palabra* aplicando un procedimiento derivado de la lingüística descriptiva en el que, aunque no se sepa estrictamente hablando qué es una palabra, se la utiliza por su "validez psicológica", como lo comenta *La lingüística. Guía alfabética* de André Martinet. Dijimos que esa "palabra" intuitiva y precientífica tenía una función de "muleta" para la descripción, *heurística* desde el punto de vista científico. Podemos decir que, en el capítulo anterior, para determinar la palabra fonológica *usamos* el significado. También introdujimos la teoría del signo de Saussure para darle sentido a ese procedimiento y poder afirmar que, para que haya signo, hace falta que toda expresión verbal tenga indisociablemente *significante* y *significado*. En este capítulo desarrollaremos sistemáticamente las buenas razones que tenemos para definir el significado como parte central de la naturaleza de la unidad *palabra* en cualquier lengua.

1. CONCEPTOS BÁSICOS

La historia de la lingüística nos permite darnos cuenta del muy largo proceso de reflexión sobre las lenguas que, desde la Antigüedad, particularmente desde la filosofía y la gramática griegas, ha venido produciéndose en las culturas que hemos heredado de Europa. Ese proceso fue diferente, por lo que muestra la historia de la lingüística, del que se produjo en India, muy ligado a la conservación de los libros religiosos del brahmanismo, los *Vedas*, orientado a la fonética, o del que se haya podido producir en China o en Mesoamérica (de estos dos grandes ámbitos culturales de la Antigüedad sabemos todavía muy poco).

La reflexión de origen griego admiraba su lengua como instrumento para llegar a conocer la naturaleza o la *esencia* de las cosas. Es decir, a pesar de que el griego se consideraba como la única lengua que les parecía verdadera y correcta (la palabra *bárbaro*, que quería decir algo como "balbuceo", o sea, expresión de quien no sabe o no puede hablar, se utilizaba para calificar a todas las demás lenguas conocidas por ellos, diferentes del griego), no se planteaba, sin embargo, como un objeto de conocimiento por sí mismo. Dicho de otra manera: la lengua era una especie de pantalla semitransparente que tenía, detrás de ella, las cosas; y eran las cosas mismas las que interesaba conocer. La gramática se concebía no como lo hace ahora la lingüística, sino como el estudio de la capacidad de las partes de la oración (sustantivo, verbo, adjetivo, etc.) para ayudar a develar aquella esencia de las cosas.

Con la aparición de los estados nacionales a partir del Renacimiento, las lenguas llamadas entonces "vulgares", para distinguirlas de las cultas latín y griego, o también "modernas", como el español, el francés, el italiano, etc., se convirtieron en interés de los gramáticos, que hasta entonces sólo se habían dedicado al estudio cuidadoso de las lenguas clásicas. Nuestro mejor ejemplo es Elio Antonio de Nebrija, un andaluz del siglo XV que se fue a estudiar a Italia con los grandes humanistas de su época y, al regresar como profesor de la Universidad de Salamanca, para enseñar latín e introducir el humanismo italiano a la universidad, llevó consigo la inquietud por conocer su propia lengua, como lo había comenzado a hacer, cien años antes, Dante en su famoso libro *De vulgari eloquentia*. Nebrija fue el primero en Europa en proponerse hacer una gramática de su lengua materna: la *Gramática de la lengua castellana*. Al hacerlo, puso en práctica dos procedimientos que habrían de resultar cruciales para el posterior desarrollo de la lingüística: el primero consiste en haber *comparado* el castellano con el latín; el segundo, en haber usado *la gramática latina como instrumento de análisis* de la estructura del castellano.

Son dos las obras que conviene leer para darse una buena idea del pensamiento griego antiguo sobre la lengua: el *Sobre la interpretación* de Aristóteles y el diálogo *Cratilo* de Platón.

A base de esos dos procedimientos, por primera vez en la historia, las lenguas se convertían, por ellas mismas, en objeto de conocimiento. Los intereses de los antiguos griegos, que venían a ser propiamente intereses filosóficos, se circunscribieron al ámbito de la filosofía (de la lógica y la ontología); los nuevos intereses de conocimiento de las lenguas modernas modificaron, en consecuencia, la comprensión de la gramática, para la que desde entonces las *partes de la oración* comenzaron a interesar como formas y como funciones, aunque todavía no por sí mismas, sino para educar a los hablantes en el manejo de una lengua culta, que no se corrompiera, como concebían la evolución del latín cuando desapareció la civilización romana. Por eso la gramática se definió por mucho tiempo como "el arte de hablar correctamente una lengua", no como un sistema interno de toda lengua, que es lo que hacemos ahora.

1.1. La objetivación, condición para la lingüística

Pero lo que aquí nos interesa es comprender cómo, poco a poco, por motivos diferentes, las lenguas fueron convirtiéndose en *objetos de estudio*. Dicho de otra manera: cómo se dio paulatinamente un proceso de *objetivación* de las lenguas. Sin ese largo proceso histórico, la lingüística como tal, es decir, como disciplina científica con objeto propio, no habría podido crearse. Fue sobre todo a partir de Ferdinand de Saussure en Europa, y del interés despertado por el antropólogo alemán emigrado a Estados Unidos de América, Franz Boas, por estudiar las lenguas de los pueblos aborígenes norteamericanos, como se formó la lingüística en que hoy nos educamos.

1.2. Actividad y producto

Hay que darnos cuenta de que la lengua se nos manifiesta de dos maneras: como una *actividad* humana, que ocurre en el tiempo cuando se habla, o en el espacio cuando se escribe, y que forma parte de muchas de nuestras acciones cotidianas; o como un *producto* codificado en grabaciones magnetofónicas o en textos escritos, dispuesto ante nuestros oídos y nuestros ojos para que lo aprendamos o lo estudiemos. Cuando nos interesamos por la lengua como actividad, como algo que *se hace* al hablar, la comprendemos como una práctica, como una *energeia*, que es el concepto —tomado del griego— con que sintetizó esa manera de

comprenderla Guillermo de Humboldt (hermano de Alejandro, el científico viajero que vino a América), un pensador alemán situado en el gozne intelectual entre la Ilustración del siglo XVIII y el romanticismo del siglo XIX. Cuando la lengua se nos presenta registrada, codificada en libros o en grabaciones, desligada de la actividad en que se produjo, la comprendemos como un *ergon*, como un resultado, como un producto, tal como nos lo propuso Humboldt.

En términos de estudio lingüístico es mucho más fácil concebir la lengua como *ergon* que como *energeia*, pues cuando se nos presenta como producto, como resultado, no sólo nos da tiempo para reflexionar sobre ella, sino que nos permite practicar conmutaciones entre diferentes signos, hacer listas de signos semejantes, compararla con otras lenguas, etc. En cambio, concebirla como *energeia* parece forzarnos a estar presentes siempre en una situación de habla, que es fugaz, que es infinitamente variada y que, por lo tanto, dificulta sobremanera nuestra necesidad de reducir el habla a sus invariantes: pasar del *habla* a la *lengua*, como lo pedía Saussure.

De ahí que la lingüística moderna, desde Saussure hasta la actualidad, pueda caracterizarse como una *lingüística del producto*, de lo ya dicho, del *ergon*. Aun más, debido al proceso de objetivación de la lengua que ha seguido la lingüística, al que debe su constitución como ciencia, la época moderna se caracteriza por dos actitudes: el interés por la lengua en sí misma —desligada de toda relación con la historia, la sociedad, la mente, la práctica—, que podemos calificar como *inmanentista*, y la tendencia a *cosificarla*, ya sea como un sistema que se contiene a sí mismo, del cual el *habla* es un accidente imperfecto, o ya sea como un órgano del cuerpo humano, como los brazos o las piernas, que se forma genéticamente en cada individuo y es capaz de funcionar con plena autonomía de la voluntad del individuo y de la sociedad en la que vive.

En este libro optaremos, al contrario, por *considerar la lengua una actividad, una energeia*, necesariamente ligada a la experiencia humana, a la práctica del hablar, al desarrollo histórico y al mundo social en que el ser humano encuentra su plenitud. Una concepción de esta clase se califica hoy en día como *pragmática*.

Una *lingüística pragmática*, en consecuencia, no niega la necesidad científica de objetivar los fenómenos que estudia, ni la necesidad de registrarlos como productos del hablar, pero se orienta por la idea de que la lengua es una actividad humana inmersa en la sociedad y la historia, ligada al resto de las prácticas de las sociedades y los individuos.

1.3. Acción y acto verbal

Decimos que *hablar es una acción*. Lo es tanto como mover una mano, correr, saludar a alguien, andar en bicicleta o leer. Cuando un niño de dos o tres meses de edad comienza a mover una sonaja para oír el ruido que hace, o cuando llora para pedirle a su madre que lo mime, ejecuta acciones. Para ese mismo niño, hablar —desde los 12 o 15 meses de edad— es otra acción, que no puede desligarse del conjunto de acciones en que acontece: grita ¡*mamá*! al mismo tiempo que extiende sus brazos hacia ella para hacerla ir a él; dice *guaguá* cuando se le aproxima un animal, o cuando lo quiere para jugar con él, o cuando oye de lejos un ladrido; articula un ¡*éste*! [éte] señalando con su dedo un juguete que le llama la atención, o esperando que alguien se lo acerque, o simplemente para dejar constancia de que lo ha percibido. Un niño vive varios años de su infancia sin darse cuenta de que el idioma que habla puede objetivarse y, por lo tanto, distinguirse de las acciones de las que forma parte. Se suele dar el caso de niños bilingües, que cambian de lengua según el interlocutor que tienen enfrente, sin poder distinguir una lengua de otra. Así, un hijo de madre alemana y padre mexicano habla con toda naturalidad alemán con su madre y español con su padre, sin saber que cambia de lengua cada vez que lo hace; incluso sin saber que *Hund* en alemán y *perro* en español, son palabras diferentes. La conciencia de que habla dos lenguas despunta años más tarde, cuando comienza a *objetivar* los signos.

Sólo la educación, que en todos los pueblos del mundo comienza en los momentos en que los padres corrigen alguna pronunciación errónea de su hijo, en que le hacen distinguir entre dos palabras, por ejemplo, entre *perro* y *gato* (cuando al niño le daba lo mismo la acción verbal con que hablaba de cualquiera de esos dos animales), o entre el alemán y el español, sólo en esos momentos comienza a objetivarse y a distinguir su lengua de sus acciones. Pero lo que hay que hacer notar es que la actividad verbal del individuo es la base de su aprendizaje de la lengua ma-

terna, y que su memoria nunca deslindará las acciones en que la aprendió, de lo que aprendió de ella.

Pero, a la vez, va aprendiendo en qué *situaciones* se habla, qué *condiciones* dan validez a sus acciones verbales e incluso qué *forma* deben adoptar esas acciones para poderse dar a entender plenamente. Imaginemos, por ejemplo, la reacción de los abuelos del niño, que llegan de visita a su casa y éste, al verlos, grita: ¡*caca*! Abuelos amorosos al fin, no le darán de inmediato un sopapo, sino que, pasmados, se apresurarán a justificarlo con "no lo regañen, no sabe lo que dice, es muy chiquito todavía". El niño acaba de aprender, inconscientemente, que esa acción verbal no corresponde a la situación de saludo y recibimiento de sus abuelos.

Modificando un poco un ejemplo de John L. Austin, un filósofo inglés precursor del punto de vista que estamos explicando, consideremos el caso de un borracho que sale en la noche de una taberna cercana a los astilleros de Southampton, y se encuentra de pronto frente a la proa de un trasatlántico que la reina de Inglaterra habrá de bautizar al día siguiente. Previsores como buenos ingleses, los empleados de los astilleros ya tenían preparados desde la noche anterior el templete para la reina y la botella de champaña amarrada a un cordón de seda, para que su majestad procediera al bautizo a la mañana siguiente. El borracho decide, en medio de la noche, los humos del alcohol y la niebla, bautizar él mismo al barco, por lo que lanza la botella al casco y dice: ¡te nombro *Diana de Gales*! Todas sus acciones correspondían a una ceremonia de botadura de un barco, sólo que no cumplían con las condiciones que hacían válida la ceremonia: sobre todo, con la de ser la reina de Inglaterra. Su acción verbal no cumplía con las *condiciones del acto*; por lo tanto, el bautizo que ejecutó el borracho no fue válido y la reina pudo llegar al día siguiente a botar el barco "como debe de ser".

Por último, consideremos ahora lo que podría ser una obra del llamado "teatro del absurdo": cuando llegan los espectadores al teatro, el telón está descorrido y en el foro, sin ningún arreglo, hay dos sillas más o menos desvencijadas. La primera reacción de los espectadores es pensar que se equivocaron de día para la función, o que se equivocaron de teatro, pues lo normal antes de una función teatral es que el telón esté corrido ante el foro y dentro de éste se encuentre una escenografía. Pero el acomodador les asegura que sí hay función, por lo que, de inmediato, comienzan a buscarle el sentido a las sillas, comienzan a buscar la *inteligibilidad* de los actos que verán en seguida, en relación con una *forma* teatral.

Lo que quieren mostrar los ejemplos anteriores es que muchas acciones pueden no cumplir con las condiciones que la sociedad les pone para que tengan sentido: pueden fallar en situación, en validez y en forma. Incluso hay veces en que acciones de esa clase resultan tan alejadas o discrepantes de las condiciones de validez que les impone la sociedad, que resultan por completo ininteligibles, ininterpretables.

Lo que requiere una acción verbal para adquirir validez es una forma socialmente definida y esa forma, determinada en cierta situación y por ciertas condiciones de validez, es lo que llamamos *acto verbal*. Un bautizo, una promesa, un saludo, la pregunta a un maestro, la exposición de un argumento ante colegas, ante un juez, ante los padres de uno, la aseveración de que algo ha sucedido, etc. son *actos verbales*, que tienen ciertas condiciones de validez para que puedan adquirir sentido para el resto de la sociedad.

Los actos verbales adquieren su sentido en cada cultura; hay muy pocos que sean actos universales. Es universal, por ejemplo, la pregunta a una persona acerca de algo que se ignora, pero no es universal, por ejemplo, el bautizo o la petición de mano de una novia.

2. LA ACCIÓN Y EL ACTO VERBAL DE DENOMINACIÓN

Una acción verbal que se encuentra en todas las lenguas y es, por lo tanto, universal, consiste en *ponerle nombre a las cosas*, es decir, en denominar objetos físicos, acontecimientos, experiencias, sensaciones, etc. Para la mayor parte de los seres humanos —por no decir que para todos—, la capacidad de denominar objetos es la característica más obvia de las lenguas. Incluso muchos pensarán que, si las lenguas no sirven para eso, son absolutamente inútiles. Es verdad que ese "ponerle nombre a las cosas" puede resultar muy sencillo o muy complicado, según el grado de precisión con que cada persona lo quiera hacer. Un poeta, por ejemplo, nos dirá que nombrar su experiencia del amor o de la nostalgia no se logra con sólo las expresiones "amor" o "nostalgia", sino que ha requerido todo el poema para acercarse con la mayor precisión posible a su propia e íntima experiencia.

Así que una acción verbal de denominación puede consistir de una sola "palabra" (entre comillas, porque seguimos todavía sin haberla definido plenamente), de una oración o de un texto relativamente largo y complejo. Sin embargo, la denominación más evidente para cualquier persona se logra con expresiones verbales de poca duración o de poca longitud, si consideramos lo escrito.

2.1. La génesis del acto de denominación

Decíamos antes que el niño pequeño no distingue sus acciones verbales de las otras acciones que ejecuta durante su vida. Cuando ha alcanzado cierta edad (generalmente alrededor del primer año de vida) y ha ejercitado su capacidad fonatoria y auditiva, dice, por primera vez en su vida, "mamá" (hay que notar que en muchas lenguas del mundo, y no sólo en español, esa primera expresión se logra con sonidos bilabiales y con vocales centrales, como lo sostuvo Roman Jakobson en su famoso estudio acerca del lenguaje infantil y la afasia, publicado primero en alemán como *Kindersprache, Aphasie und allgemeine Lautgesetze*, y posteriormente en inglés, con modificaciones, con el título de *Studies on Child Language and Aphasia*). Como lo ha explicado convincentemente el psicólogo suizo Jean Piaget, esa expresión forma parte de un conjunto complejo de acciones infantiles que, en ese momento, el niño no puede separar. Así, la expresión "mamá" no es todavía la denominación de su madre, sino parte de un conjunto de acciones que el niño ha ido realizando ya por cierto tiempo, entre las cuales está la audición de la expresión "mamá". Por ejemplo, la experiencia de la protección, la ternura, la alimentación, la cercanía, el olor, el tacto de la piel de su madre, y la de necesidad de cuidado y de alimento.

Así que el primer "mamá" del niño no comienza por ser un acto de denominación, sino una manifestación del complejo de acciones del que forma parte *conjuntamente*. Pero esa combinación de acciones ha ido dando lugar a lo que, según Piaget, es el proceso central de la formación de la inteligencia del ser humano: por un lado, a la creación de *esquemas de conocimiento senso-motor* (puesto que el niño no sólo percibe, sino también actúa) que dan lugar a invariantes, es decir, a *formas* de conocimiento y, por el otro lado, a la paulatina distinción de elementos "interesantes" del complejo de acciones. Uno de esos elementos "interesantes" llega a ser la articulación de la voz "mamá". Pasará todavía cierto tiempo antes de que el niño *disocie* la voz "mamá" no sólo del complejo de acciones, sino también de la presencia de su madre. Cuando eso suceda, cuando pueda articular "mamá" fuera del complejo de acciones en que aprendió y desarrolló la voz, y la utilice para llamarla aunque esté completamente satisfecho en alimento, en cariño, etc. y su madre esté ausente es cuando podemos considerar que se produce una acción de denominación de su madre. Según el estudio de estos procesos, que constituye el núcleo de la teoría de Piaget, esto sucede en el quinto estadio de la formación de la inteligencia, hacia los 18 meses de edad.

(Por eso suele suceder, en las familias, que al niño parezca darle lo mismo durante cierto tiempo si al pronunciar "mamá" está jugando con su abuela, tiene entre sus manos un biberón o ve entrar a su habitación a cualquier mujer, con lo que su madre se enoja o se pone celosa.)

Pero hay que notar que, aunque la voz *mamá* esté afincada en la capacidad fono-auditiva del niño —lo que es un hecho físico—, lo que la vuelve signo lingüístico de la lengua materna del niño es el esfuerzo que éste realiza por ir aproximando su expresión a la que oye de sus mayores. Es decir, es una acción denominadora, pero dirigida por una forma definida por la lengua de los adultos que lo rodean. El niño no nombra arbitraria y caprichosamente a su madre, sino que se ajusta a una forma socialmente fijada.

El acto de denominación no se produce, en consecuencia, arbitraria y azarosamente; no proviene de una pura voluntad del individuo ni de una convención que él mismo impone. Lewis Carroll, el famoso autor de *Alicia en el país de las maravillas*, compuso una escena en el libro *A través del espejo, lo que Alicia encontró al otro lado*, que muestra graciosamente esta idea:

Dice Humpty-Dumpty (que en la versión española del libro recibió el nombre de Zanco Panco) a Alicia: "... lo que demuestra es que hay 364 días para recibir regalos de incumpleaños.

—Desde luego— asintió Alicia.

—¡Y sólo uno para regalos de cumpleaños! ¡Ya ves, te has cubierto de gloria!

—No sé qué es lo que quiere decir con eso de la 'gloria'— observó Alicia.

Zanco Panco sonrió despectivamente.

—Pues claro que no..., y no lo sabrás hasta que te lo diga yo. Quiere decir que 'ahí te he dado con un argumento que te ha dejado bien aplastada'.

—Pero 'gloria' no significa 'un argumento que deja bien aplastado', objetó Alicia.

—Cuando *yo* uso una palabra— insistió Zanco Panco con un tono de voz más bien desdeñoso— quiere decir lo que yo quiero que diga, ni más, ni menos".

Alicia se opone a él, diciéndole que ese no es un derecho suyo. Un niño no podría utilizar una voz a la manera en que lo hace Humpty-Dumpty; más bien, acepta que las voces tienen una forma y un significado, que le vienen dadas por su lengua materna (aunque hay un pequeño ámbito en que sus padres aceptan que el niño nombre objetos inventando palabras, pero generalmente esos casos son muy reducidos, como los hipocorísticos, o "nombres de cariño" que se dan a los abuelos o los hermanos).

La denominación, en consecuencia, no procede de simples acciones denominadoras, sino de *actos* que tienen ciertas condiciones de validez, de las que hay que destacar sobre todo dos: que se trata de acontecimientos sociales y que proceden de una historicidad de la lengua materna que, para todo individuo, está previamente determinada.

Un acto de denominación se puede dar en situaciones como las siguientes:

- Un niño ve por primera vez un león y pregunta a sus padres: "¿qué es eso?", bajo la suposición de que sus padres saben lo que es y le van a decir verdaderamente lo que es; no que lo van a engañar. Su padre responde: "es un león" o "se llama león" o "un león".
- Una persona explica alguna cosa y nota que su auditorio no ha entendido una de sus expresiones. Se interrumpe y aclara: "por 'hipocorístico' entiendo un nombre de cariño que se pone a una persona".
- Un sacerdote, frente a la pila del bautismo, dice: "yo te nombro *Juan*".
- Un científico descubre cierta partícula atómica nunca antes teorizada o percibida y dice: "esta partícula se llamará *quark*".

Todos estos actos nos parecen aceptables, en cuanto a las situaciones en que se dan, las condiciones de validez que conocemos en nuestras comunidades lingüísticas y las formas adoptadas para realizarlos. Son *actos de denominación*, tal como los hemos definido.

2.2. Las expresiones verbales del acto de denominación

En los ejemplos anteriores, las expresiones verbales de la denominación han sido varias: "ese es un león", "se llama león", "un león", "hipocorístico", "yo te nombro Juan", "esta partícula se llamará *quark*". Pero si nos acercamos sobre todo a las ciencias y las técnicas contemporáneas, veremos que varios otros tipos de expresiones verbales pueden aparecer en dichos actos. Por ejemplo, el frenesí armamentista ha producido denominaciones como "proyectil aire-aire", "proyectil balístico intercontinental" y varios más; la contabilidad nos da "producto interno bruto" o "impuesto al valor agregado"; la medicina, "síndrome de inmunodeficiencia adquirida", etc. En náhuatl clásico, sus especialistas nos explican que "in xochitl in cuicatl" es el nombre de la poesía, y se dice que en alemán es posible (aunque no del todo real) la siguiente expresión:

"Donaudâmpferschifffahrtsgesellschaftskapitânwitwenrentenauszahlungstag"

(que quiere decir: "día de pago de las pensiones de las viudas de los capitanes de la Compañía de Vapores del Danubio").

La existencia de estas unidades de denominación indica que, aunque toda "palabra" es una unidad de esta clase, no todas las unidades de esta clase son palabras, sino que pueden tener una forma más compleja o más larga. De donde se puede concluir que, *si bien un acto de denomi-*

Es verdad que se cometen errores de interpretación en las situaciones en que se producen estos actos. Se cuenta, por ejemplo, que la palabra *canguro*, que el español tomó del inglés *kangaroo*, por ser esta la lengua de los primeros exploradores occidentales de Australia, no era la denominación aborigen de ese animal —como lo supusieron los exploradores ingleses— sino que quería decir en alguna lengua australiana "no te entiendo". Nos podemos imaginar muy bien la escena: un explorador inglés señala al canguro y espera la respuesta de su informante; pero éste no entiende lo que quiere señalar el explorador y le contesta, lógicamente, *kangaroo* 'no te entiendo'; el inglés asume que esa forma es el nombre aborigen del canguro. (No hay datos que permitan afirmar que ese es el origen de la palabra *kangaroo* en inglés —y de allí, en el resto de las lenguas no australianas—, por lo que puede tratarse de un caso de lo que se llama "etimología popular", como se verá más adelante en este libro.)

nación que produce una unidad de denominación es una condición necesaria para que haya palabra, tampoco es una condición suficiente para ello.

Sin embargo, ahora podemos comprender por qué esa "intuición" de la palabra, que se manifiesta en la denominación de objetos y experiencias, resulta tan útil a la lingüística descriptiva y "psicológicamente válida" para tantos lingüistas: el descriptor de una lengua desconocida, que se enfrenta a ella por primera vez, no tiene otra manera de reunir un número suficiente de signos de esa lengua para comenzar su análisis, que acudiendo a la ostensión de objetos físicos presentes, cuya denominación espera recibir de su informante. Aquello que le permite esperar esa denominación, y no un gran discurso, es la universalidad del acto de denominación y el hecho de que las "palabras" son, ante todo, nombres.

3. PROCEDIMIENTOS DE OBTENCIÓN DE UNIDADES DE DENOMINACIÓN

El reconocimiento de unidades de denominación no ha formado parte, hasta ahora, del arsenal de procedimientos de investigación de la lingüística descriptiva, por el motivo explicado en el primer capítulo: la inexistencia de una teoría del signo para los creadores de esa lingüística; se han basado en ese reconocimiento, como se ha venido insistiendo, como apoyo heurístico de la descripción, pero no le han dado el lugar teórico que le corresponde en la construcción de la ciencia del lenguaje. Por eso no lo considerarían un nivel de la descripción.

En este *Curso*, en cambio, en que se exponen los motivos y los argumentos para reconocer a la unidad de denominación un papel fundador de toda la descripción, se puede considerar ese reconocimiento, a la vez, como constitutivo de un *tercer nivel de análisis* lingüístico, del que se derivan varios procedimientos de trabajo.

3.1. Preguntas acerca de objetos físicos mostrados

Dado que el acto verbal de denominación es universal, se pueden derivar de ese hecho varios procedimientos que permiten al lexicólogo obtener unidades de denominación. El más común es aquel que consiste

en enseñar objetos físicos sencillos, existentes en la cultura material de la comunidad lingüística considerada. Es necesario aprovechar toda la situación en que se producen esos actos. Por lo general, cuando dos personas se encuentran por primera vez, o por primera vez en un día, se saludan o se presentan. Ambos actos dan lugar a unidades de denominación, como "buenos días", "hola", "*Grüss Gott*" (algo como "que Dios esté contigo" en Alemania meridional) o "Pedro Martínez", "su servidor, el alcalde", etc. Todo lingüista descriptivo toma nota inmediatamente de esas expresiones, para comenzar a recabar signos necesarios para el análisis. Puesto que el lingüista hace lo mismo, en su propia lengua, hacia su informante, se establece una situación social concreta, de comienzo de diálogo. En vez de quedarse callado el lingüista, como si fuera un juez, un testigo o un psicoanalista, el siguiente paso puede ser comenzar a mostrar objetos, nombrándolos en su propia lengua, para invitar al informante a ofrecer equivalentes en la suya. Se sigue la hipótesis de que esas respuestas ofrecen unidades de denominación en la lengua considerada.

Tenemos un caso semejante en los siguientes ejemplos, tomados de una entrevista a una niñita huichol:[1] se le muestran varios animales y se le ha ido preguntando cada vez: "*¿'iki ketitita*? '¿éste qué es?'" y la niña responde "*kauxai tsiriki* 'pues es una zorra'", en que *kauxai* se aísla como unidad de denominación de la zorra; a "*¿'iki ta*? '¿y éste?'", responde "*miki ta maye* 'ese es un león'" (*maye*); "*¿ketitita 'iki*? – *erepante*" (un elefante, con el hispanismo).

3.2. Unidades de cita

La experiencia de lingüistas descriptivos dedicados a la lexicografía de lenguas amerindias nos provee de otro procedimiento. Éste se basa en el hecho, comprobable en cualquier lengua, de que todo hablante reflexiona acerca de su propia lengua, aunque no lo haga con los instrumentos que ha desarrollado la lingüística. Así por ejemplo, cuando el hablante está relatando algo, puede detenerse para explicar una de sus "palabras"

[1] Agradezco los ejemplos a la profesora Paula Gómez, de la Universidad de Guadalajara, Jalisco.

que merece aclaración, del mismo modo que quien lo oye, puede preguntarle algo como "¿qué quiere decir X?" En esos casos, se separa una unidad de denominación, que además tiene la particularidad de que convierte la "palabra" en cuestión en una forma útil para citarla, forma a la que habrá que dedicar mucha atención en los dos siguientes capítulos. En náhuatl moderno, de la localidad de Ocotal Texizapan, Veracruz, por ejemplo, [motalowa] 'corre', [nehnemi'] 'camina', [pata:ni'] 'vuela', son formas de cita de unidades de denominación, que posteriormente se reconocerán como verbos; en tanto que los sustantivos se citan, o bien sin morfemas de número ni género: [ikpál] 'silla' (en los diccionarios *icpalli*), [ilwika'] 'cielo' (*ilhuicatl*), o bien con un morfema "posesivo" [noma] 'mi mano', [nohyish] 'mi ojo', [nohyia] 'mi nariz', cuya función, más que "posesiva" es la de actualizar el nombre.[2]

Concluimos, pues, en que la unidad de denominación tiene un papel teórico central en la explicación del fenómeno real de la palabra, aunque no baste su reconocimiento para definir suficientemente lo que es una palabra en cualquier lengua que se considere. A la vez, podemos utilizar esta idea para explicar el papel heurístico que ha tenido la palabra en la descripción lingüística, y asignarle un nivel propio de descripción —el tercero— en el procedimiento de descubrimiento de los signos de una lengua, para lo cual hemos ofrecido dos métodos de trabajo: el de las respuestas a preguntas de identificación de objetos físicos y el de las unidades reflexivas de cita.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

actividad (*energeia*); producto (*ergon*); acción y acto verbal; situación; condiciones y forma en el acto verbal; unidad de denominación; unidad de cita.

[2] Debo estos ejemplos a mi alumna Rubí Ceballos.

OBRAS CONSULTADAS

AUSTIN, John L., *Cómo hacer cosas con palabras. Palabras y acciones*, compilación de J.O. Urmson, traducción de G.R. Carrió y E.A. Rabossi, Barcelona, Paidós, 1982.

BUHLER, Karl, *Teoría del lenguaje*, Madrid, Revista de Occidente, 1967 (1a. ed. alemana, Jena, 1934).

CARROLL, Lewis, *A través del espejo, lo que Alicia encontró al otro lado*, traducción de Jaime de Ojeda, Madrid, Alianza Editorial, 1992.

COSERIU, Eugenio, *Teoría del lenguaje y lingüística general. Cinco estudios*, Madrid, Gredos, 1973.

HUMBOLDT, Wilhelm von, *Über die Verschiedenheit des menschlichen Sprachbaues und ihrem Einfluss auf die geistige Entwickelung des Menschengeschlechts*, Berlín, 1836.

JAKOBSON, Roman, *Studies on Child Language and Aphasia*, La Haya, Mouton, 1971.

PIAGET, Jean, *La formación de la inteligencia en el niño*, Madrid, Aguilar, 1969.

3
LA PALABRA MORFOLÓGICA

INTRODUCCIÓN

En los capítulos anteriores tratamos dos condiciones necesarias de la determinación de la palabra, correspondientes a lo que ahora podríamos considerar como un primer acercamiento a su significante y su significado: la de su complexión fonológica y la de su valor denominativo. Estudiamos también los varios procedimientos de que disponemos para definir inicialmente esos dos planos del signo, independientemente de la lengua que se esté investigando. Pero como se trata de condiciones que, por sí mismas, no bastan para determinar correcta y adecuadamente lo que es una palabra, hay que continuar explorando otros de sus elementos constitutivos. En este capítulo nos ocuparemos de la morfología de la palabra.

1. CONCEPTOS BÁSICOS

En el primer capítulo (1, § 1.2) introdujimos la teoría del signo lingüístico que construyó Ferdinand de Saussure, e hicimos alusión a una expansión explicativa de esa teoría que elaboró Louis Hjelmslev. Ahora desarrollaremos esta última expansión para que se puedan situar mejor los niveles de descripción, que ya hemos explicado, y hacer evidente la complejidad fenoménica de la unidad *palabra*.

1.1. Forma y sustancia; contenido y expresión

Es relativamente fácil entender cómo, del continuo sonoro que pueden producir los órganos fonatorios del ser humano, sólo ciertos sonidos pueden convertirse en fonemas de una lengua. Así por ejemplo, una ca-

rraspera no da lugar a fonemas (al menos, no parece haberse descubierto una lengua en que esa posibilidad tenga lugar); el hipo, que se produce por algún desarreglo estomacal, tampoco da lugar a fonemas; un timbre de voz muy agudo no se convierte en fenómeno suprasegmental de una lengua, sino que, en el mejor de los casos, le da cierto carácter a la pronunciación de una comunidad lingüística, además de ayudar a reconocer la voz de una persona. La gran variedad de sonidos que pueden producir los órganos fonatorios humanos, pero que todavía no se convierten en sonidos "de una lengua", es lo que, siguiendo a Hjelmslev, podríamos llamar "materia sonora" y no es objeto de la lingüística por ella misma, en su mera realidad acústica. Esos sonidos se llegan a convertir en sonidos de una lengua (y en objeto de la lingüística) cuando *se conforman* en unidades perceptualmente discretas de una lengua histórica particular.

La investigación de la *percepción* de los sonidos de una lengua es todavía una materia pendiente de la lingüística, que ha preferido explorar su carácter acústico, que es un hecho físico (la fonética experimental), no de la percepción, y los medios de que disponemos para formular las observaciones realizadas (diferentes escuelas de la fonología formal contemporánea). El fenómeno de la percepción de los sonidos de una lengua no ha llamado la atención, hasta ahora, ni a la lingüística ni a la psicología (tanto a la de la percepción, como a la de la forma). Pero ese faltante de la lingüística y de la psicología no nos impide considerar, de manera suficientemente verosímil, que nuestra percepción es la que segmenta el continuo sonoro en unidades discretas, que después podremos reconocer como fonemas. Tales unidades discretas son resultado de una *forma* que les da la lengua histórica.

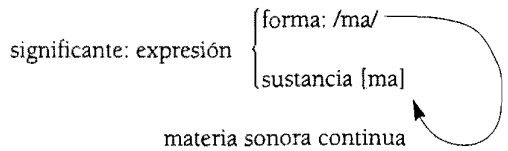
Sostener que el sonido de una lengua es un efecto de una forma nos crea, sin embargo, una dificultad importante, ya que podría pensarse que la forma es anterior al sonido, cuando, desde un acercamiento inocente y factual al fenómeno, parecería ser secundaria en relación con el sonido mismo; por otra parte, postular una anterioridad de la forma nos conduciría a ciertas posturas llamadas "idealistas" de la filosofía o "innatistas" de la lingüística, según las cuales esa forma antecede siempre a la lengua y, si no es parte de las llamadas "potencias del alma" en el cristianismo, entonces debe ser una impronta genética contenida en el ácido desoxirribonucleico (ADN) humano. Dicho más fácilmente: parece lógico plantearnos la pregunta de qué es primero: el sonido o su forma.

Para decidirlo y eliminar los riesgos de "idealismo" e "innatismo", volvamos a la observación de lo que sucede con los niños de pocas semanas de edad. Uno puede darse cuenta de que un bebé empieza muy pronto a emitir sonidos con su boca: balbuceos, chasquidos de la lengua, ronroneos, sollozos, silabeos, etc. Durante las semanas que el bebé pasa haciendo estos ejercicios fonatorio/perceptuales, ninguna de las personas que lo rodean dirán que el bebé "comienza a hablar"; sólo está produciendo materia sonora, aunque, como lo explica Piaget, está profundamente enzarzado en la construcción de esquemas de asimilación y adaptación del sonido a su inteligencia. Pero basta con que cualquier día alguien sorprenda en su fonación una sílaba como [ma], para que de inmediato crea "reconocer" una sílaba de su lengua materna y se ponga a animar al bebé para que la repita, repitiéndosela a su vez muchas veces y estimulándolo para que la imite. Esa sílaba se ha vuelto importante porque corresponde a una *forma* de la lengua *de quien la oye*. Es, en consecuencia, la persona que oye la fonación del bebé quien le asigna valor a ciertos sonidos, los vuelve pertinentes para la lengua y, con ello, *les da forma*, para poderlos considerar elementos de la lengua que habla. Dicho de otra manera: *la forma de la lengua histórica*, que en efecto antecede al aprendizaje de la lengua de cada persona, es la que selecciona materia sonora para volverla sonido de la lengua.

En ese sentido es como hay que interpretar la relación forma/sonido: el sonido adquiere pertinencia verbal para el bebé, porque quienes lo rodean, lo destacan y le dan valor; no han aceptado como elementos sonoros de la lengua cualquier emisión de voz, sino sólo la que corresponde a las formas de su lengua materna. La forma selecciona o determina el valor del sonido; lo precede históricamente, no ideal ni genéticamente.

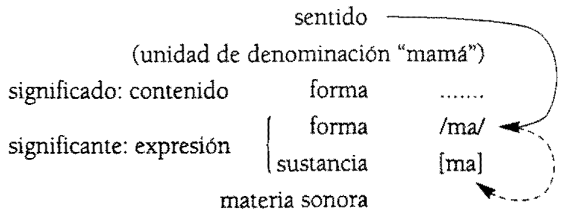
A la vez, en cuanto la forma es la que determina las características de las invariantes que constituyen la fonología de la lengua, aquellos sonidos concretos que emite el bebé y valen como emisiones de su lengua materna, se convierten en la sustancia real de sus fonemas: es decir, en la *fonética* de la lengua de ese bebé.

Se puede esquematizar este proceso de la siguiente manera:

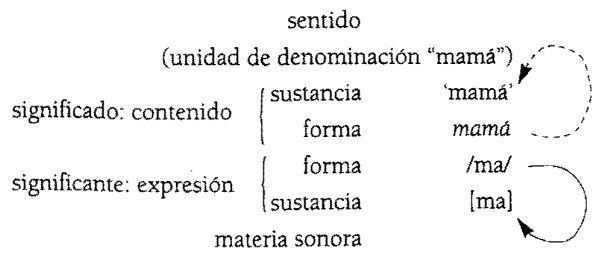


Al mismo tiempo, el sonido percibido como *sustancia de la expresión* de una lengua se ha reconocido como tal gracias al significado que le reconoce el oyente. Aquella sílaba [ma] que destacó la persona que oía al bebé, se volvió pertinente porque podía interpretarse como un primer ensayo de la voz /mamá/; si no fuera así, se podría considerar que, en su desarrollo articulatorio, el bebé alcanza una fase de silabeo, es decir, de ejercicio fonatorio/perceptual de secuencias consonánticas y vocálicas, que no llega a convertirse en emisión verbal y sigue siendo materia sonora.

Así que fue la unidad de denominación de la lengua adulta —mamá— la que dio valor a aquella sílaba y la convirtió en forma de la expresión del signo, primero, y después en su sustancia. Ahora podemos ampliar el esquema anterior de la siguiente manera:



Decíamos en el segundo capítulo que las unidades de denominación nombran algo, elaborando aquello que pueda "tener sentido" para una persona. Nombrar a la madre tiene un sentido, que se manifiesta en la forma adulta "mamá", y a partir de ese momento se produce una forma del contenido determinada, que es la palabra *mamá*. Podemos, entonces, completar el esquema anterior y llegar a la concepción completa del signo saussureano, tal como la explicaba Hjelmslev:



Si ya se ha explicado que la disciplina lingüística que estudia la sustancia de la expresión es la fonética, y la que estudia la forma de la expresión es la fonología, ahora podemos concluir diciendo que el estudio de la sustancia del contenido corresponde a la semántica, en sentido estrecho, y el de la forma del contenido a la morfología y a la sintaxis. Hay razones para considerar que el estudio completo del plano del contenido, y no sólo de su sustancia, es el dominio de la semántica; sin embargo, para hacernos entender por la mayor parte de las corrientes lingüísticas, haremos esa distinción entre semántica, de un lado, y morfología y sintaxis (o morfosintaxis), de otro (sin que eso implique que aceptemos que la semántica, la morfología y la sintaxis sean compartimentos o subsistemas separados de una lengua, que obliguen a preguntarse en qué consiste su diferencia, por un lado, y su interrelación, por el otro —lo que hoy en día han dado en llamar "la interfaz semántica-sintaxis").

1.2. Morfología y sintaxis

La morfología y la sintaxis son las partes de la lingüística que se ocupan del estudio de la forma del contenido de los signos lingüísticos. Por lo general, se considera que la morfología se dedica a estudiar la constitución formal de la palabra, en tanto que la sintaxis trata de las relaciones que se producen entre las palabras en una oración. Como se ve, ambas disciplinas lingüísticas presuponen la existencia de las unidades *palabra*, que aquí venimos tratando de reconocer. Puesto que, a estas alturas, aunque ya sabemos que la palabra tiene dos características determinantes (su estructura fonológica y su valor como unidad de denominación), todavía no hemos terminado de definirla por completo, haríamos mal en basarnos en

esas definiciones de la morfología y la sintaxis, si ellas a su vez dependen de la definición de la unidad *palabra*. Se crearía un círculo vicioso.

1.3. La teoría de la doble articulación

Para resolver este problema acudamos a otro principio general de la definición de lo que son las lenguas, explicado por el lingüista francés André Martinet: la llamada "teoría de la doble articulación".

Ya habíamos visto en el primer capítulo que el número de fonemas que constituye una lengua no es muy grande, sino que oscila entre 20 y 40, algunos más o algunos menos. También que el número posible de combinaciones que se pueden hacer entre estos fonemas, de dos en dos, de tres en tres, etc., hasta llegar, si se quiere, a combinaciones de 12 en 12 o más, da una cifra astronómica.

Si los signos lingüísticos sólo consistieran de combinaciones de fonemas, la cifra de unidades "palabra" que se podrían crear en cada lengua sería tan alta, que resultaría inútil para los propósitos de comunicación entre seres humanos que definen la existencia de las lenguas. ¿Qué memoria humana podría recordarlas?

También dijimos en el primer capítulo que esa cantidad de combinaciones posibles de fonemas quedaba restringida por las estructuras posibles de las sílabas, por la diferente repartición de los fonemas en una cadena, por las posibles repeticiones de fonemas, etc., pero, aunque así se reduciría la cifra obtenida, seguiría siendo enorme.

Por el contrario, todos los seres humanos nos podemos dar cuenta de que las combinaciones de fonemas que constituyen nuestras palabras dependen de otra clase de arreglos de su forma, que no son arreglos fonológicos, sino que éstos se supeditan a otras clases de unidades formales, cuyos esquemas no son cuantiosos, pero que nos permiten significar todo lo que queramos. Esos esquemas están determinados por el significado.

Observando, entonces, que una lengua se articula por elementos que *portan* significado (raíces, desinencias, palabras, oraciones) y elementos que *materializan* ese significado (fonemas), Martinet propuso que una de las características centrales de toda lengua humana es su *doble articulación*: los elementos que portan significado son unidades de primera articulación; los que sirven solamente para materializar, es

decir, a hacer posible la transmisión del significado, sin ser ellos mismos significativos, sino sólo distintivos, son unidades de segunda articulación. Dicho de otra manera: los elementos de primera articulación son *significativos*; los de segunda articulación, *distintivos*. Como se ve, la distinción de Martinet ofrece una perspectiva diferente del signo lingüístico, en comparación con la perspectiva saussureana; la de Martinet está determinada no por la naturaleza del signo lingüístico como significado y significante, sino por el papel que desempeñan los fonemas (unidades de segunda articulación) y los morfemas, palabras y esquemas oracionales (unidades de primera articulación) en la economía y el rendimiento informativo de los sistemas lingüísticos, pues gracias a las características de las unidades de primera articulación se puede limitar de manera importante el número de combinaciones posibles de fonemas en una cadena; y gracias a la cantidad de fonemas y esquemas silábicos de una lengua se pueden construir varios cientos o miles de unidades de primera articulación, que, sin embargo, son menos que las que podría componer una simple combinación de fonemas.

De acuerdo con esta teoría, la forma del contenido de los signos lingüísticos corresponde del todo a la primera articulación; la unidad *palabra* tiene su lugar en ella. Así podemos hacer un primer reconocimiento de la forma del contenido, sin tener que distinguir todavía entre palabras y oraciones, o entre morfología y sintaxis.

2. EL MORFEMA

En el primer capítulo (1, § 1.3) explicamos el funcionamiento de la conmutación para poder distinguir fonemas (invariantes) de pronunciaciones diferentes (variantes) de los mismos fonemas. Es decir, utilizamos ese método para reconocer unidades de la segunda articulación. El método de la conmutación también se aplica para reconocer unidades de la primera articulación. Así por ejemplo, en la oposición entre:

depósito/ deposito/ depositó

que habíamos utilizado para reconocer la función distintiva del acento en español, podemos ahora mirar las diferencias de significado a que da

lugar la posición del acento (una muestra más de la *reciprocidad* del significante y el significado). En esta oposición descubrimos que esas tres formas tienen una *base de la comparación*, que es el segmento que permanece igual al hacer las conmutaciones: /deposito/, y tres elementos diferentes, dados por la posición del acento, cuyos significados son: *a*] en *depósito*, la posición esdrújula del acento sirve a significar el sustantivo 'lugar en donde se coloca alguna cosa para guardarla, almacenarla o acumularla'; *b*] en *deposito*, la posición grave del acento sirve a significar el verbo conjugado en primera persona del singular del presente de indicativo, 'acción de colocar o poner alguna cosa en cierto lugar para guardarla, almacenarla o acumularla'; *c*] en *depositó*, la posición aguda del acento sirve a significar el mismo verbo anterior, conjugado en tercera persona del singular del pretérito de indicativo.

Si ahora, siguiendo el método de *pares mínimos* explicado en 1, § 1.3. oponemos:

deposito : depositó
canto : cantó
bailo : bailó
toco : tocó
miro : miró

podremos observar que todas estas formas dan lugar a cinco bases de las comparaciones:

deposit-, cant-, bail-, toc- y mir-

y a dos elementos que sirven para diferenciar las formas presentes de los verbos de sus formas pretéritas: *-o*, *-ó*. Correspondientemente, si definimos como bases de las comparaciones los segmentos *-o* y *-ó*, diferenciamos cinco segmentos ligados con ellas.

Lo mismo podemos hacer con las siguientes unidades: *niño/niña/niñito/niñita/niñero/niñera*, *perro/perra/perrito/perrita/perrero/perrera*, en que sus respectivos significados distinguen como bases de las comparaciones: *niñ-*, *perr-*, y el género, masculino o femenino, el diminutivo, la ocupación (*perrero*, *niñera*), la afición (*niñero*) y el lugar en donde habitan los perros (*perrera*); de donde obtenemos:

niñ-, perr- y -o, -a, -it-, -er-

Todos estos resultados nos ofrecen unidades de primera articulación, reconocidas por su valor significativo.

Como hispanohablantes, que además hemos ido a la escuela elemental, nos ha resultado fácil darnos cuenta de que la segmentación de las unidades de denominación consideradas tenía que dar lugar a esas formas y no a otras. Pero lo normal para un lingüista es enfrentarse con lenguas que desconoce, y por lo tanto, podría ensayar otras conmutaciones, para buscar posibles segmentaciones diferentes de las mismas unidades, por ejemplo:

ni-ño o ni-ñ-o, pe-rro o pe-rr-o

Para poder segmentar *ni-ño* o *pe-rro* habrá que buscar otras unidades de denominación que contengan alguno de los cuatro segmentos propuestos, cuya oposición dé lugar a pares mínimos, que descubran un rasgo significativo entre ellos.

Para ello, pongamos en juego otras unidades de denominación. Por ejemplo, *niño/nido/nilo* o *perro/peto/pelo*. Para que *ni-* y *pe-* fueran unidades de primera articulación, los significados de esas seis unidades de denominación deberían depender de ellas. Pero descubrimos que no es *ni-* lo que contribuye a reconocer el significado de *niño*, *nido* y *nilo*, del mismo modo en que *niñ-* lo hizo en *niña*, *niñero*, *niñera*; *ni-* es una sílaba que forma parte de varios cientos de unidades de denominación en español, de significados muy diferentes entre sí, que no forman, por así decirlo, "parte de una familia"; lo mismo podemos decir de *pe-* en *perro*, *peto*, *pelo*. Algo semejante ocurre si se ensayan las segmentaciones *pe-t-o*, *pe-l-o* y *ni-d-o*, *ni-l-o*.

Para comprobar que este método de investigación nos da resultados, en cualquiera que sea la lengua en estudio, tomemos ahora algunos ejemplos del tzeltal de Oxchuc:

yašmo'on / yašmo'at / yašmo' / yašmo'tik

Sin saber todavía qué significan, por la pura forma de estas cuatro unidades de denominación podemos proponer como base de la compa-

ración *yašmo'* y distinguir *-on*, *-at*, *-tik*. Nos preguntaremos entonces qué resultados obtendríamos con otra base de la comparación en relación con las tres formas con que finalizan esas unidades de denominación:

ya'šwe'on / yašwe'at / yašwe' / yašwe'otik

y concluiríamos en que se puede segmentar una base *yašwe'-* y los mismos finales de unidad que habíamos aislado antes, aunque con una pequeña variante: *-otik*.

Nuestro informante nos dijo que los significados de esas unidades son, aproximadamente, 'subo, subes, sube, subimos' y 'como, comes, come, comimos', por lo que podemos concluir que *-on*, *-at* y *-(o)tik*, más la falta de un signo para la tercera persona del singular (que muchos lingüistas representan con 0) son elementos de la conjugación de esos verbos.

Con los mismos procedimientos llegaríamos a la conclusión de que *yaš-* añade al significado de esas expresiones un aspecto imperfectivo.

Podemos así comprobar que los resultados de la conmutación basados en las diferencias de significado, que dieron lugar a esas segmentaciones, nos ofrecen segmentos mínimos, que ya no se pueden seguir analizando sin dejar de ser significativos. Aquellos fragmentos de signo obtenidos como segmentos mínimos significativos mediante una descomposición dirigida por el significado son *morfemas*.

Por eso se ha definido siempre el morfema como una unidad mínima con significado, pero hay que aclarar que *el morfema no "tiene significado" por sí mismo, sino en conexión con otro morfema, con el que forma las unidades de denominación a las que significan*, pues si se considerara que cada morfema tiene significado en sí mismo, la consecuencia sería que el significado de la unidad de denominación sería el resultado de una composición de los significados de los morfemas que la constituyeran, lo cual es falso, de acuerdo con el papel determinante que se ha dado a las unidades de denominación: el significado 'animal carnívoro cuadrúpedo que se caracteriza por mover la cola cuando algo le llama la atención o está contento' no es resultado de un supuesto significado de *perr-* compuesto o yuxtapuesto con el de *-o*. *La segmentación de las unidades de denominación no descubre una composición de sus significados, sino que solamente reconoce unidades mínimas de valor significativo: los morfemas.*

2.1. Morfemas libres y ligados

Si el morfema es la unidad mínima de primera articulación, lo que se quiere decir es que se trata de la más pequeña unidad significativa que pueda resultar de una segmentación de una expresión verbal. Así, *comer* se puede segmentar en *com-er*; *comeríamos* en *com-er-ía-mos*; *niño* en *niñ-o*; *mesa* en *mes-a*; *mosca* en *mosc-a*; mientras que *árbol*, *mujer*, *azul*, *hoy*, *yo*, *peor* no pueden descomponerse más. Todos estos elementos son morfemas; sólo que *com-*, *-er*, *niñ-*, *-o*, *mes-*, *-a*, *mosc-*, *-a* requieren unirse con otros para poder reconocer suficientemente su significado y no pueden aparecer en aislamiento (salvo, por supuesto, en un texto de lingüística o de gramática, en donde, sin embargo, señalamos su dependencia con un guión), en tanto que *árbol*, *mujer*, *azul*, *hoy*, *yo*, *peor* no pueden fragmentarse más y no necesitan de otros morfemas para manifestarse. A los primeros los llamaremos *morfemas ligados* y a los segundos, *morfemas libres*.

Estas unidades de denominación formadas por morfemas libres, que además constan de un número entero de sílabas y están compuestas por fonemas que pueden tener funciones demarcativas, ya son *palabras*. Es decir, hay morfemas libres que son, a la vez, *palabras*. Llamemos a esta clase *palabras tipo 1*.

Podemos complementar esta *primera clase de unidades palabra* con una prueba común en la lingüística descriptiva, que consiste en hacer preguntas a sus informantes en las que aparezcan estos morfemas como respuestas únicas:

—¿Cuándo vendrá su hijo?
—*hoy*.
—¿Cómo sigue el enfermo?
—*peor*.
—¿De qué color es el cielo?
—*azul*.
—¿Quién hizo esto?
—*yo*.

¡Cuidado! Hay que hacer estas preguntas sin inducir la clase de respuestas que se buscan. Por ejemplo, no es válido hacer una pregunta como ésta: —Su casa está en la acera de en... —*frente*. Pues en esta pregunta, el interrogador está, o bien presuponiendo que la palabra es *enfrente*, o que la respuesta debe estar constituida por una preposición *en* y un sustantivo *frente*.

2.2. Paradigmas de morfemas

En los ejemplos anteriores del español (pero podríamos hacer lo mismo con el tzeltal) pudimos notar que hay dos clases de morfemas ligados: unos, como *perr-*, *niñ-*, *nid-*, *pet-*, *pel-*, cuya cantidad es prácticamente ilimitada en una lengua, y otros como *-o*, *-ó*, *-ero*, *-er*, *-ía-*, *-mos*, que vemos aparecer constantemente en las unidades de denominación que ponemos en juego para llevar a cabo las conmutaciones, y cuyo número es mucho menor que el de los primeros. Martinet designa a los morfemas del primer tipo como *morfemas de inventario abierto*; a los del segundo, como *morfemas de inventario cerrado* (nuevamente, véase que se trata de una caracterización solamente cuantitativa). Unos y otros pueden sustituirse por otros de su mismo tipo en una cadena de morfemas. Por ejemplo:

Dada la cadena *niñ-o*, *niñ-* puede sustituirse por muchísimos morfemas que sirvan para construir sustantivos, como *perr-*, *pet-*, *pel-*, y podemos agregar *banc-*, *herman-*, y cientos más; en la misma cadena, *-o* puede sustituirse, sin embargo, sólo por: *-a*, *-it-*, *-er-*, *-ez*, *-erí-*, *-ic-* (niña, niñito, niñero, niñez, niñería, niñico, etc.) y quizá unos pocos más, pues se trata de morfemas de inventario cerrado; *-a* es, en español, un morfema de la flexión nominal, en tanto que los demás son morfemas derivativos.

Los morfemas de inventario abierto reciben el nombre de *lexemas* (las *raíces* de la terminología gramatical tradicional). Todas las lenguas tienen una gran variedad de lexemas, que podríamos poner en una larguísima lista. Esa lista podría considerarse un gran *paradigma*, pero a la vez subdividirse en cuantos paradigmas pareciera conveniente, según las propiedades formales o de sustancia de contenido que encontráramos en sus miembros. Así por ejemplo, el paradigma de los verbos de la primera conjugación en español, el paradigma de los sustantivos absolutos terminados en *-tl* del náhuatl, el paradigma de los lexemas que construyen unidades de denominación de la habitación, como *cas-*a, *edific-*io, *palac-*io, *choz-*a, *departament-*o, etcétera.

Los morfemas de inventario cerrado han recibido el nombre de *gramemas* en las obras de muchos lingüistas, a partir de la enseñanza de Martinet. Las listas de gramemas de una sola clase, por ejemplo, los de cada tiempo verbal en español, como *-o*, *as*, *-a*, *-amos*, *-áis*, *-an* del presente de indicativo, forman también *paradigmas*.

Entendemos por *paradigma* un conjunto de unidades lingüísticas, ya sea de segunda articulación (fonemas) o de primera (morfemas y, más tarde, palabras) que se caracterizan por oponerse unos con otros y ser capaces de sustituirse mutuamente en una cadena de segmentos determinada.

3. MORFOTÁCTICA

Como bien sabemos, toda expresión verbal se desarrolla en el tiempo. Esta es una característica de todas las lenguas. Cuando hablamos, unas unidades verbales salen de nuestra boca antes que otras; cuando escribimos, cualquiera que sea el sistema de escritura, alfabético, silábico, ideográfico, unos símbolos se escriben antes que otros. Al distinguir entre morfemas libres y ligados se ha manifestado con claridad esa propiedad de las lenguas. Puesto que, en muchos casos, esos morfemas son elementos de paradigmas (como, en los ejemplos anteriores, los de tiempo verbal en español), podemos comprender que la secuencia de morfemas sea a la vez una secuencia de paradigmas de morfemas.

En cada lengua, según el tipo de que se trate, sus secuencias morfemáticas tienen cierto orden, más o menos rígido. Así por ejemplo, en español la manifestación del tiempo verbal sólo puede suceder al lexema (aparece como *sufijo*) y no antecederlo (no aparece como *prefijo*); en cambio, los pronombres que significan los complementos del verbo, pueden anteceder al lexema y al verbo mismo, como en *me-lo-das* (posición *proclítica*), o sucederlo, como en *dá-me-lo* (posición *enclítica*).

Llamamos *morfotáctica* a la característica que tienen todas las expresiones verbales de una lengua, de ordenar sus morfemas en una secuencia determinada o en varios esquemas secuenciales.

Una de las tareas más complejas del estudio morfológico de las palabras de una lengua es definir con claridad sus esquemas morfotácticos. Cada lengua tiene cierta cantidad de esquemas morfotácticos, un fenómeno semejante a la variedad de esquemas silábicos, a los que

aludimos en el primer capítulo de este libro, o a la variedad de los tamaños de la palabra medida en sílabas. Aquí trataremos de definir los criterios principales para determinar esquemas morfotácticos de cada lengua orientados, por supuesto, al reconocimiento morfológico de la unidad palabra.

Dada una unidad de denominación, los morfemas que la constituyen aparecen en una secuencia determinada. Para encontrar el esquema morfotáctico de una unidad de denominación y decidir si se trata de una unidad *palabra* hace falta: *a*] segmentar la secuencia en morfemas, con los criterios antes expuestos; *b*] probar la cohesión que haya entre ellos, y *c*] determinar el orden en que aparecen.

3.1. Segmentación

Es bien claro que toda emisión verbal, de la longitud que sea, puede descomponerse en morfemas libres y ligados, eludiendo la existencia probable de unidades *palabra*. Es eso lo que suelen hacer muchos lingüistas, para quienes la palabra no tiene existencia real. Tomemos el siguiente ejemplo:

> "La primavera es una clara ducha
> Que nos lava de sombra y telarañas".
>
> TOMÁS SEGOVIA
> "Canción de libertad en primavera"[1]

A manera de experimento, dejaremos en suspenso el reconocimiento de unidades de denominación; y obviando todas las conmutaciones que habría que hacer si este texto fuera de una lengua desconocida, podemos segmentarlo como sigue:

la-primavera-es-un-a-clar-a-ducha-que-nos-lav-a-de-sombr-a-y-telaraña-s.

En él podemos reconocer morfemas libres que son palabras: *primavera* y *ducha* (como se las definió arriba en el § 2.1), *es* (al que nos vol-

[1] Del libro *Salir con vida. Poemas 2000-2002*, México, Ediciones Sin Nombre, 2003.

veremos a referir en el capítulo 6) y morfemas claramente ligados: *un-*, *-a*, *clar-*, *-a*, *lav-*, *-a*, *sombr-*, *-a*, *telaraña-*, *-s*.

De esta segmentación podemos obtener lexemas (*primavera*, *lucha* y *es* son, a su vez, *palabras*; además: *clar-*, *lav-*, *sombr-* y *telaraña-*) y gramemas: *la*, *un-*, *-a*, *que*, *nos*, *de*, *-s*. (Recordemos que la distinción entre lexemas y gramemas depende exclusivamente de su membresía en paradigmas abiertos o cerrados.) Además hay elementos de cuya libertad o ligazón no estamos seguros, sobre todo porque en nuestro sistema de escritura se consideran "palabras": *y*, *la*, *de*, *que*, *nos*, aunque requieren otras unidades para construir un significado: *la*, en este ejemplo, necesita relacionarse con un sustantivo femenino; *de* con dos sustantivos, uno a cada lado, o con un verbo y un sustantivo; *que* e *y*, con oraciones, y *nos* con un verbo. El problema de si estos morfemas son palabras también lo trataremos más adelante.

3.2. Cohesión

Por ahora, es posible observar que los morfemas que obtenemos del análisis de *clara* (*clar-a*), de *lava* (*lav-a*), de *sombra* (*sombr-a*) y de *telarañas* (*telarañ-a-s*) —en que *-a* forma parte, o bien del paradigma de morfemas de género (*-a*, *-o*: *claro/clara*), o bien del paradigma de morfemas de tiempo presente de indicativo (*lava/lavo/lavas*, etc.)— tienen con sus respectivos lexemas una *cohesión*, es decir, quedan estrechamente ligados, como si entre ellos hubiera una especie de pegamento o "glutinosidad". Podemos detectar la cohesión si intentamos intercalar algún morfema perteneciente a otro paradigma (algunos lingüistas llaman a este procedimiento "prueba de intercalabilidad"). Por ejemplo, en *clar-a*, *sombr-a* y *telarañ-a* no podemos intercalar ningún otro morfema de flexión nominal, aunque sí un morfema de diminutivo *-it-* que da *clarita*, *sombrita*, *telarañita*, o en otras variedades del español *-ic-* para obtener *clarica*, *sombrica*, *telarañica*; también se pueden intercalar morfemas derivativos como -uch- (*clarucha*), pero no muchos más. Entre los lexemas y los morfemas de tiempo en español, sólo se intercala la llamada "vocal temática" *-a-*, *-e-*, *-i-* (*am-a-r* frente a *am-e*, *com-e-r* frente a *com-ía* y *com-í*, *sub-i-r* frente a *sub-ía* y *sub-í*). Se puede considerar, entonces, que la *cohesión* entre morfemas varía entre un máximo —los casos de la flexión

nominal y del lexema o raíz verbal en español y la vocal temática—, algunos grados intermedios (el caso de la derivación) y ninguna cohesión: los morfemas que aparecen en contacto, aunque ligados, dejan total libertad para intercalar otros morfemas entre ellos).

Cuando hay cohesión máxima entre dos morfemas sucesivos consideramos, siguiendo al lingüista estadounidense Joshua Greenberg, que se forma un *núcleo morfemático*. Así, *ama-* (el lexema *am-* y la vocal temática *-a-*) en *amaba*, *amaría*, *amaré*, etc. o, en el caso de la flexión nominal, *niñ-o*, *niñ-a*. La cantidad de núcleos que se puede encontrar en una secuencia de morfemas es siempre menor o igual que el número de morfemas obtenidos.

Si se representa con la letra **m** el número de morfemas y con la letra **n** el número de núcleos en una unidad de denominación, diremos que:

$$m \geq n$$

Para poder verificar el procedimiento de segmentación morfológica seguido y la prueba de cohesión entre morfemas que hemos ensayado en la lengua española, tomemos ahora por caso la siguiente expresión en alemán, una lengua caracterizada como "aglutinante" en la vieja tipología de las lenguas que hemos heredado en nuestra educación:

diegeschäftsordnungdesreichstagesannehmen

('aceptar el reglamento interno del parlamento')

Obviando, como antes, todas las conmutaciones entre probables elementos de primera articulación que hacen falta para una segmentación morfológica de la expresión (que es la tarea más común para un lingüista que quiére estudiar una lengua que desconoce), llegamos al siguiente resultado:

die-ge-schäft-s-ordn-ung-des-reich-s-tag-es-an-nehm-en

Entre *die-* y *ge-* pueden intercalarse muchos morfemas e incluso cadenas de morfemas, como *die* gross-e *ge-*, *die* heut-ig-e *ge-*, *die* seit Jahren nicht mehr gültige *ge-*, etc., por lo que no hay cohesión entre ambos

morfemas y, por lo tanto, no forman un núcleo morfemático. Entre *ge-* y *-schäft-* no puede intercalarse ningún otro morfema, por lo que forman entre sí un núcleo morfemático *geschäft-*. Entre *geschäft-* y *-s-* puede intercalarse otro morfema, como *-igkeit-*, para dar *geschäftigkeits-*, pero no parece haber más posibilidades, por lo que su cohesión es, por lo tanto, alta. Entre *geschäfts-* y *-ordn-* pueden intercalarse algunos morfemas como *-zu-*: *geschäftszuordn-*, o como *-an-*: *geschäfts-an-ordn-*, por lo que la cohesión entre ambos es débil y no forman núcleo entre sí. Entre *ordn-* y *-ung-* no pueden intercalarse otros morfemas, por lo tanto, hay cohesión máxima entre ellos: *ordnung-*.

Hasta aquí hemos podido distinguir claramente dos núcleos morfemáticos en alemán: *geschäft* y *ordnung*. Si representamos la falta de cohesión con el símbolo #, la cohesión mediana con el símbolo / y la cohesión máxima con el guión (-) tendremos, hasta ahora:

#die#ge-schäft/s/ordn-ung

Entre *ordnung-* y *-des-* pueden intercalarse muchos otros morfemas y, en consecuencia, no se forma núcleo morfemático; entre *des-* y *-reich-* lo mismo; entre *reich-* y *-s-* sucede lo mismo que entre *geschäft-* y *-s-*, por ejemplo: *reich-tum-s*; entre *reichs-* y *-tag-*, encontramos el mismo fenómeno que entre *geschäfts-* y *-ordn-*; entre *tag-* y *-es* no pueden intercalarse otros morfemas, por lo que se forma el núcleo *tages*, de donde descubrimos un núcleo morfemático más (aunque, como veremos más adelante, hay otra posibilidad que no consideraremos ahora, para no complicar más el ejemplo); el paso de *tages-* hacia *-an* permite multitud de morfemas entre ellos, por lo que no forman un núcleo; entre *-an-* y *-nehm-* pueden intercalarse algunos morfemas, cuya cohesión es mediana. En cambio, entre *nehm-* y *-en* no puede intercalarse nada, de donde descubrimos un núcleo morfemático más. Así, en la expresión analizada se han encontrado cinco núcleos: *geschäft-*, *-ordnung*, *reich-*, *tages* y *-nehmen*. El análisis nos da:

#die#ge-schäft/s/ordn-ung#des#reich/s/tag-es#an/nehm-en#

Cuando ponemos en juego el significado, como habíamos dicho, buscamos unidades de denominación. Un hablante de alemán nos dirá

que *geschäft*, *ordnung*, *reich*, *tages* y *nehmen* son unidades de denominación; sus significados son, respectivamente, 'asunto, negocio', 'orden', 'reino, dominio', 'del día' y 'tomar'.

3.2.1. Capacidad de sustitución paradigmática

Si los resultados de la segmentación morfológica nos han llevado a proponer la existencia de cinco núcleos morfemáticos en el ejemplo anterior, y sabemos que sus significados son unidades de denominación, podemos ahora analizar una más de las propiedades que solemos atribuir a nuestra idea tradicional de la palabra: el que sea ampliamente sustituible por unidades de la misma clase, dado un sintagma. Para ello, podemos hacer una prueba más, que consiste en conmutar los núcleos morfemáticos obtenidos con otros que también sean unidades de denominación y formen parte de inventarios abiertos:

Podemos comprobar que *geschäft* puede conmutarse con *büro* 'oficina, despacho', *firma* 'empresa', *tätigkeit* 'actividad', etc.; *ordnung* con *reihe* 'serie', *liste* 'lista', etc.; *reich* 'reino' se puede conmutar con *land* 'país o tierra', o con *republik*; *tag* con *nacht* 'noche', *periode* 'período', *jahr* 'año', etc.; *nehmen* se conmuta con *ablehnen* 'rechazar' o *vertagen* 'suspender', *geben* 'dar', *bringen* 'traer', etc. Comprobado lo anterior, ya podemos reunir todas las *condiciones morfológicas necesarias* para decidir que hemos descubierto cierta clase de unidades *palabra*. Además, tomando en cuenta las condiciones necesarias previas de complexión fonológica y valor denominativo, podemos afirmar que, las tres juntas, sólo juntas, se vuelven *necesarias y suficientes*.

Esos núcleos morfemáticos del alemán, reconocidos por su cohesión interna, y que corresponden a unidades de denominación, en la medida en que tienen un número entero de sílabas y posibles fonemas con función demarcativa, *son palabras*. Lo mismo podemos afirmar de los núcleos morfemáticos encontrados en español: *clara*, *sombra* y *telaraña*. Las llamaremos *palabras tipo II*.

3.2.2. Núcleos morfemáticos con formas supletivas

En cambio, entre los núcleos morfemáticos descubiertos en los ejemplos del español, *ama-* todavía no se corresponde con una unidad de denominación. Por lo tanto, hay que proceder a analizar la morfotáctica de expresiones como *amaba*, *amaría*, *amé*, etc. Se verá que hay cohesión máxima entre el núcleo *ama-* y los morfemas temporales *-ba*, *-ría*, etc., pero que la hay también entre *am-* y *-é*, en donde se pierde la vocal temática. Este es un fenómeno muy común en todas las lenguas: debido a las características de la evolución histórica que da lugar a todas las lenguas habladas, aparecen, a los ojos del lingüista interesado por encontrar en ellas una sistemática regular y sencilla, *anomalías* entre sus morfemas, que se hacen más evidentes cuando se comprueba que, pese a sus diferencias, forman parte de los mismos paradigmas.

Como el objetivo de este *Curso* no es el estudio pormenorizado de los diferentes procesos morfológicos evolutivos que dan lugar a las palabras de la lengua histórica (objetivo de un curso sobre formación de palabras, tanto en su sistemática como en su historia), sino exclusivamente el reconocimiento de la unidad *palabra*, podemos, por un lado, sostener que se puede dar cuenta de secuencias como *ama-ba*, *ama-ría*, *am-é* con los mismos instrumentos conceptuales que hemos venido utilizando; y por el otro, que la descripción del paradigma tanto del lexema correspondiente como de los gramemas ligados a él en cohesión máxima, obliga a tratarlos como matrices, que proveen de elementos a las correspondientes reglas de formación de las palabras, lo que tomaremos en cuenta en el capítulo 6.

Así, la matriz del paradigma del lexema de *amar* es:

[am-, ama-]
(diferentes lingüistas llaman a esta clase de paradigmas, "polimórficos"; a las formas que los componen, *supletivas*)

y la matriz del paradigma de lexemas del tiempo verbal del presente de indicativo es:

[-o, -as, -a, -amos, -áis, -an].

De donde se puede sostener que, si bien un núcleo morfemático no necesariamente es equivalente con una palabra (salvo en las arriba ejemplificadas de tipos I y II), en casos como el de los verbos en español, en que pueden proponerse paradigmas con formas supletivas para unificar sus núcleos morfemáticos, tenemos palabras del tipo II.

3.2.3. Secuencias de núcleos morfemáticos

Pero, volviendo a nuestro ejemplo del alemán, todavía no agotado, debemos ahora agregar que el mismo hablante que nos dijo que *geschäft*, *ordnung*, *reich*, *tages* y *nehmen* son unidades de denominación (y, en consecuencia, palabras del tipo II), nos dirá que *geschäftsordnung* y *reichstages* son, a su vez, unidades de denominación que significan, respectivamente, la primera, 'reglamento interno', un concepto muy común de las organizaciones sociales y la administración. La segunda, *reichstages*, quiere decir 'del parlamento' y en este caso, ningún hablante de alemán reconocerá en el segmento *tages* alguna relación con el 'día'. Es decir, no se puede analizar *reichstages* como una secuencia de dos palabras, cuyos significados sean 'reino' y 'del día', que es lo que hicimos antes, conmutando morfemas aislados, sin tomar en cuenta el significado de toda la cadena; así que morfológicamente se pueden reconocer dos núcleos morfemáticos en *geschäftsordnung* y en *reichstages*, pero sus significados no corresponden a la secuencia de dos palabras, sino a nuevas unidades de denominación.

Si ponemos en práctica la prueba de capacidad de sustitución paradigmática con otras unidades de denominación, como hicimos en el parágrafo anterior, resulta que *geschäftsordnung* puede conmutarse con *geschäftsreise* 'viaje de negocios', *geschäftsmann* 'hombre de negocios', *benutzungsordnung* 'reglamento de uso', e incluso con *tagesordnung* 'orden del día', etc.; *reichstages* puede conmutarse con *Parlament* o *Landtag* 'parlamento de un Estado federal'.

Hay lenguas, como el alemán, el náhuatl y muchas más, que tienen una gran capacidad para crear unidades de denominación a base de la secuencia de dos o más núcleos morfemáticos, organizados alrededor de sendos lexemas. El ejemplo más extremo, morfológicamente posible, pero no del todo real, es el que encontramos en el segundo capítulo de este *Curso* (§ 2.2):

Donaudampfschifffahrtsgesellschaftskapitänwitwenrentenauszahlungstag

Si separamos en esta unidad de denominación exclusivamente los núcleos morfemáticos tendremos:

Donau/dampf/schiff/fahrt(s)/gesellschaft(s)/kapitän/witwe(n)/
rente(n)/(aus)zahlung(s)/tag

Es decir, 10 núcleos en una secuencia particular, unidos por los morfemas *-s-* y *-n-* con que se manifiesta el caso genitivo alemán, cuya función consiste en especificar la clase de objetos significados por el núcleo morfemático más a la derecha. En este ejemplo, *tag* 'día' se especifica con *auszahlung* 'pago': día de pago; *auszahlungstag* se especifica con *renten* 'pensiones': día de pago de pensiones, etcétera:

'día de pago de las pensiones de las viudas de los capitanes
de la Compañía de Vapores del Danubio'

Este ejemplo, considerado por cualquier hablante de alemán como una exageración juguetona de las posibilidades morfotácticas de la palabra alemana, nos sirve para mostrar dos características más de las unidades de denominación en las lenguas: en primer lugar, que las reglas de formación de palabras, desde el punto de vista exclusivamente morfológico, pueden dar lugar a expresiones de esta clase, de manera indefinida (lo que demuestra que la constitución morfológica de la palabra es una condición necesaria para su determinación, pero no suficiente por sí sola); en segundo, que incluso una lengua como la alemana limita sus posibilidades de formación de palabras mediante su significado y que éste, como se verá en el capítulo 4, se compone a base de esquemas de conocimiento que no actúan como meras composiciones de elementos.

Consideremos en seguida otros problemas que nos plantean las secuencias de núcleos morfemáticos. A un hispanohablante no se le habrá escapado que *telaraña* podría tener alguna similitud con los ejemplos del alemán que hemos venido analizando, sobre todo si es fácil reconocer en ella que se trata de una *tela* de una *araña*, es decir, que ambos núcleos son unidades de denominación tan libres, que pueden reconstituir

el significado de *telaraña* por separado, como lo hemos hecho, en tanto que en *reichstages*, al separar los dos núcleos, se destruye el significado de la unidad de denominación. Unidades como *telaraña*, *sacacorchos*, *matamoscas*, etc., obtienen su significado de los significados de los dos núcleos que las forman. Podemos, por lo tanto, confirmar la conclusión expuesta arriba de que el análisis en secuencias de núcleos morfemáticos por sí solo, no es condición suficiente para reconocer la existencia de unidades palabra. Cuando dos núcleos se yuxtaponen para formar una unidad de denominación, es el significado de la unidad morfológica resultante el criterio determinante para decidir si se trata de *palabras* o de *composiciones de palabras*.

En cada lengua puede haber ambos procedimientos de formación de unidades de denominación. En alemán, por ejemplo, *aktiengesellschaft* es una composición a base de *aktien* 'acciones bursátiles' y *gesellschaft* 'compañía' (en español, su equivalencia es *sociedad anónima*; ambas son términos técnicos de las finanzas); incluso *flughafen* 'aeropuerto' puede considerarse todavía una composición de *flug* 'vuelo' y *hafen* 'puerto' (compárese con el español *puerto aéreo*). En cambio, en *handschuh* 'guante' o en *fingerhut* 'dedal', ningún hablante de alemán considerará que sean composiciones de *hand* 'mano' y *schuh* 'zapato', o de *finger* 'dedo' y *hut* 'sombrero', aunque puesto a reflexionar, sonría por la metáfora. En español, *mondadientes* 'palillo', *cortaplumas* 'navaja de bolsillo' o incluso *librepensador* 'individuo que afirma su libertad de criterio sobre todo en relación con imposiciones eclesiásticas' tampoco pueden considerarse meras composiciones, por cuanto su unidad proviene de un significado propio.

Son estas características de las secuencias de núcleos morfemáticos las que dan lugar a una permanente indecisión entre los hablantes de una lengua y entre las agencias normativas que dirigen su reflexión y su acción lingüística, como las "academias de la lengua" y los diccionarios, acerca de si se trata de palabras o de composiciones de palabras de carácter fraseológico. Considérese el largo debate a que dio lugar en la lexicografía francesa la unidad *pomme de terre* (así escrita) 'papa', que no es resultado de los significados de los núcleos *pomme* 'manzana' y *terre* 'tierra' (la intervención de la escritura se tratará en el capítulo 5).

Por eso no hay identidad entre la unidad de denominación y la palabra. Toda palabra es una unidad de denominación, pero no toda unidad de denominación es una palabra. Se pueden considerar palabras, según todo lo anterior, a las unidades de denominación constituidas por dos (o más) núcleos morfemáticos que den lugar a un significado propio, no compuesto por los significados aislados de los dos núcleos-palabra. Serán *palabras compuestas, tipo III*. En cambio, las unidades de denominación compuestas por dos o más núcleos morfemáticos cuyo significado resulte de la composición de los significados de los núcleos-palabra tomados aisladamente, no son palabras, sino sintagmas relativamente fijos que, en algún momento de la historia de la lengua, pueden llegar a convertirse en palabras compuestas; así, *máquina de coser* es un sintagma fijo.

El estudio de esta clase de unidades de denominación, que presenta gran complejidad, es materia de la fraseología.

3.2.4. Dos riesgos de los procedimientos de segmentación

Pero habíamos notado en nuestra primera segmentación del ejemplo en alemán, que *geschäft* podía analizarse en *ge-* y *-schäft*, guiándonos por una serie de conmutaciones con otras unidades de denominación como *gebirge* 'sierra, cadena de montañas', *gesicht* 'rostro, cara', *geselle* 'soltero', e incluso con los participios pasivos regulares *gemacht* 'hecho', *gesehen* 'visto', *gekommen* 'venido', etc. en los que un morfema *ge-* contribuye a construir el significado 'realizado, concluido' (algunos lingüistas lo denominarían "resultativo").

Es el significado de la unidad de denominación el que impide la segmentación de *ge-* en las tres primeras palabras, y fuerza a diferenciarlo del *ge-* resultativo en los participios pasivos regulares. Ningún hablante de alemán reconocerá hoy un morfema *ge-* colectivizador en la formación de esos sustantivos, aunque esa sea su etimología y siga siendo productivo para formar palabras denigratorias: *gesinge* aproximadamente 'tonadilla', *gemache* 'obrilla', *geschaffe* 'engendro'.

Para comprender esta diferencia, consideremos el mismo fenómeno en español: podríamos segmentar un morfema *re-* en *reforma* y *revista*, *reformar* y *revisar*, a partir del hecho de que existe un morfema *re-* del español, que da al significado de las unidades en que interviene, el valor de repetición o vuelta: lo *rehecho* es lo que se volvió a hacer; un presidente *reelegido* es el que se vuelve a elegir, un libro *releído* es el que se

volvió a leer, etc.; en consecuencia *reformar* y *reforma* serían "volver a formar (y su resultado)", *revisar* y *revista* serían ¿volver a "visar" (y su resultado)? Ningún hablante actual del español, incluso de los bien educados, reconocería espontáneamente en esas unidades la presencia del morfema *re-*. (Agreguemos, por mor de exhaustividad, que *revisar* y *revista* son préstamos del francés, que no se formaron con ese morfema *re-* del español; sí, del francés.) De la misma manera, si un hablante de alemán ya no reconoce un morfema *ge-* en *geschäft* o en *gebirge*, *gesicht*, *geselle*, mientras que sí lo hace en *gemacht*, *gesehen*, *gekommen*, entonces *geschäft*, *gebirge*, *gesicht*, *geselle*, son morfemas libres únicos.

El error de segmentar en demasía la secuencia de morfemas en una expresión verbal, cuando el lingüista no domina la lengua que estudia (lo que es el caso normal), es uno de los más comunes, en los que suele caer incluso el lingüista más avezado. Para reducir ese riesgo es necesario reunir un buen conjunto de expresiones en la lengua que se estudia, sobre todo en forma de discursos o de textos, para practicar sobre ellas un primer análisis, y posponer la interrogación directa a los hablantes hasta que se cuente con un buen número de hipótesis de segmentación, que ellos sólo confirmen o reprueben. Forzar a sus informantes a responder a un verdadero bombardeo de pruebas de conmutación entre segmentos probables los lleva, en pocos minutos, a aceptar como morfemas elementos segmentados que no lo son y a perder por completo de vista su realidad verbal.

Parte de los mismos riesgos es un proceso que podríamos llamar de "etimologización del análisis", que consiste en retrotraer las secuencias de morfemas a la constitución histórica de las palabras, incluso a formaciones sistemáticas posibles ("virtuales", dirían los estructuralistas) que nunca existieron, como se demuestra con el caso de *revista* y *revisar*, y podríamos agregar muchos ejemplos más, como *inclusión*, *intervenir*, *importunar*, *remitir*, *omitir*, que no deben analizarse como secuencias de *in-clusión*, *inter-venir*, *im-portunar*, *re-mitir* y *o-mitir*, en comparación con *inexacto* o *imposible*, en las que los hablantes reconocen la presencia del prefijo *in-*. Lenguas como el alemán y el náhuatl se prestan para esa clase de análisis erróneo, precisamente por su capacidad para construir secuencias morfológicas muy largas y con varios núcleos en ellas. Tratándose del náhuatl, en México, en donde es la lengua privilegiada por la ideología nacional, es eso lo que anima a la creación de multitud de seudopalabras nahuas, que no son sino frases descriptivas y forzadas a partir de una necesidad de traducción del español.

Algunos movimientos reivindicatorios del náhuatl en México, esparcidos más o menos por el centro del país, suelen construir esas seudopalabras nahuas sobre la base de su capacidad morfológica "virtual". Así, si en el mundo náhuatl no existían tiendas especializadas en la venta de globos, serpentinas, silbatos, platos y vasos de papel para fiestas (como es lógico pensarlo), en una tienda dedicada a esos objetos, en Tepoztlán, Morelos, fabricaron su nombre: *tlanamacalolihuipilcalli*, que no es una unidad de denominación, no es una palabra náhuatl, sino una descripción de lo que se vende allí.

4. MORFEMAS LIGADOS DE INVENTARIOS CERRADOS, CON POCA O NINGUNA COHESIÓN

Las dos oraciones que nos han servido para explicar la segmentación y la cohesión morfológica todavía tienen varios morfemas ligados que hay que considerar en relación con las características morfológicas de la palabra. Pues es claro que las expresiones *la*, ***una***, *que*, *nos*, *de* e *y*, en el ejemplo del español, no son unidades de denominación, aunque tienen una estructura fonológica de sílabas completas. Sólo pueden aparecer aislados cuando los *mencionamos* —no cuando los *usamos*— en una acción reflexiva sobre la propia lengua, como en "***la*** es un artículo determinado", o como lo estamos haciendo *para* ***hablar*** *de ellos*. En el ejemplo del alemán, *die*, *-s-*, *des*, *an-*, están en la misma situación.

Si efectuamos pruebas de conmutación entre morfemas que puedan ocupar las mismas posiciones que tienen en las oraciones ejemplificadas, veremos que todos ellos, sin excepción, forman parte de inventarios cerrados: *la* y *una* del inventario de artículos; ***nos***, del inventario de pronombres personales; *que* e *y* del de conjunciones, y *de*, del de preposiciones (el conocimiento que tenemos de nuestra gramática nos facilita la enunciación y la denominación de esos inventarios; si se tratara de morfemas de una lengua desconocida para nosotros, probablemente habríamos hecho clasificaciones más neutras de los paradigmas de los que forman parte, sin utilizar la terminología gramatical latina).

Al hacer lo mismo con los morfemas alemanes, encontraremos que forman parte de los inventarios cerrados del artículo determinado (*die*, *des*), del caso genitivo (*-s-*, *-n-*, *-es*) y de las preposiciones (*an*). Todos

esos morfemas están ligados a núcleos morfemáticos, por lo que podría tratarse de elementos de la unidad palabra, pero ya hemos visto que o no tienen cohesión con ellos o la tienen débil. Además, las unidades de denominación a las que se ligan no los requieren para poderse enunciar: se puede decir *primavera* sin anteponerle *la*; *clara ducha*, sin anteponerles *una*; *geschäftsordnung*, sin anteponerle *die*; *reichstages*, sin anteponerle *des*, etc. Lo mismo, *lava* sin posponerle *de*, etc.

Resulta que, aun cuando son morfemas ligados, pertenecientes a paradigmas cerrados, ni son unidades de denominación ni forman parte de ninguna de ellas. Se dirá que, entonces, no son palabras, aunque tengan una estructura fonológica de sílabas enteras y posibles fonemas con función demarcativa.

El caso de *annehmen* es particular, pues en una serie de conmutaciones encontraríamos que *annehmen* es una nueva unidad de denominación, cuyo significado es 'aceptar', a diferencia de *nehmen* 'tomar'; es decir, que la preposición *an* se integra, con cohesión débil, a la palabra *nehmen* y produce una nueva unidad, con significado propio. Lo que contribuye a significar esa preposición es la cercanía o el movimiento hacia el enunciador, una propiedad significativa de las preposiciones alemanas.

Hemos obtenido tres buenos resultados (las palabras tipos I, II y III) y otros insatisfactorios y sorprendentes, por lo que no es de extrañar que históricamente la lingüística se haya sentido incómoda con la noción tradicional de palabra; pues sostiene que, o todos los elementos de una oración son palabras, o la noción de *palabra* es de validez limitada.

4.1. Las palabras gramaticales

Por lo general, los hablantes de cualquier lengua se dan cuenta, aunque sea de manera oscura e imprecisa, de que esos morfemas ligados a palabras de los tipos I, II o III tienen una cohesión débil con ellas (o ninguna), no corresponden a unidades de denominación y forman parte de conjuntos reducidos de formas, que desempeñan los mismos papeles o muy parecidos en la construcción de expresiones de su lengua.

De ahí que puedan, por una parte, segmentarse siempre de la misma manera y, por la otra, conformar paradigmas cuyos miembros se

pueden enlistar sin grandes dificultades. Eso confiere a estos morfemas, como los artículos, las preposiciones y las conjunciones (o los clasificadores de muchas lenguas amerindias) una identidad propia, que resalta cuando aparecen en una secuencia de morfemas. Es su poca o ninguna cohesión con su entorno morfológico y su pertenencia a inventarios cerrados lo que los individualiza, tanto desde el punto de vista de la forma del contenido como desde el punto de vista de su presencia perceptual. Es decir: puesto que la palabra tiene una configuración que se hace patente a la percepción humana (gracias a su estructura fonológica, en que los fonemas con función demarcativa, el acento de intensidad y sus patrones silábicos, de los que hablamos en el primer capítulo, tienen un papel preponderante), los hablantes pueden separar de ella aquellos morfemas poco cohesionados y de inventario cerrado. El efecto de esas dos características es económico: es más eficaz distinguirlos y separarlos de las secuencias morfemáticas que constituyen las palabras, que unirlos a ellas, aumentando la longitud de la secuencia.

Así por ejemplo, en español o en alemán, en donde el artículo no sólo actualiza al nombre, sino que agrega información redundante al reconocimiento de su género y número, por ejemplo en: *la* cabe*z*a, *los* automóvile*s*, *una* muñeca, *die* Sonne 'el sol', *der* Mond 'la luna', *das* Mädchen 'la joven', etc., por más que tengan cohesión débil con la palabra que los sigue y deban ocupar esa posición en la secuencia morfológica, resulta más económico para distinguir la unidad de denominación que los sigue si se los segmenta y se los aísla por completo, que si se los agrega a la unidad de denominación. (Cuando uno aprende estas lenguas, es necesario aprender el artículo que corresponde a cada palabra, pues hace falta para la construcción gramatical, lo cual es prueba de su necesidad morfológica en la posición correspondiente.) Cuando se trata de clasificadores, en muchas lenguas se prefijan a la unidad de denominación y en la que, realmente, hacen falta para que su significado se entienda plenamente, el mismo hablante se da cuenta de que se trata, a pesar de ello, de morfemas que conviene aislar y considerar por separado, para reconocer la eficacia con que operan en la morfología.

En el caso de las conjunciones y las preposiciones, ya habíamos visto antes que, aunque ligadas, no tenían cohesión con los morfemas que las precedían o las sucedían, y su determinación debía proceder de otro nivel de estructuración, diferente del que da lugar a la unidad

palabra. Pero sucede lo mismo con ellas: es más económico aislarlas y considerarlas unidades aparte de todas las que constituyen las unidades palabra.

Resulta entonces que, si en el nivel en el cual se reconoce la unidad *palabra* (que en relación con los dos capítulos anteriores, podríamos considerar el *cuarto nivel de análisis lingüístico*, una vez despejadas todas las unidades *palabra*), lo que resta, se ajusta a la caracterización recién ofrecida, se extiende a esos morfemas la caracterización de *palabra*.

Es lo que han reconocido muchos gramáticos a lo largo de la historia: desde los de la Edad Media, que llamaban a esta clase de morfemas *sincategoremas* (en donde el prefijo sin-, que encontramos en palabras como *sincronización*, *síntaxis*, etc. construye el significado 'simultaneidad'), hasta la gramática actual, que unas veces los llama "palabras gramaticales" y otras "palabras relacionales".

Podemos, en consecuencia, considerarlos *palabras tipo* IV, de acuerdo con lo que la tradición de pensamiento acerca de las lenguas ha hecho. No se trata, como ligeramente lo han calificado algunos lingüistas, de capricho intuitivo, sino de un efecto de la percepción de los hablantes, concebido como principio de economía de la clasificación de los elementos de su lengua.

5. ORDEN DE MORFEMAS

El ejemplo de *annehmen* nos sirve también para ilustrar la complejidad morfotáctica que puede tener una palabra. Si en español, por ejemplo, no se pueden intercalar morfemas de valor sintáctico en la palabra, sino sólo ciertos paradigmas flexionales o derivativos, en alemán es posible intercalar en *annehmen* un morfema que vuelve infinitiva la oración en que aparece este verbo: *anzunehmen*.

De ahí la necesidad de definir con precisión tanto el orden en que pueden aparecer los morfemas en una secuencia como su grado de obligatoriedad. Ya puede uno irse imaginando que cada orden de morfemas en una palabra corresponde a un esquema formal determinado y esos esquemas son varios en cada lengua, por lo que se pueden clasificar en una lista, estipulando cuándo ciertos morfemas se prefijan al lexema, cuándo se sufijan, cuándo se intercalan. El resultado será una matriz de

esquemas morfotácticos para definir las diferentes formas de las unidades palabra en cada lengua.

Así por ejemplo, en español el patrón mínimo del verbo es raíz verbal+persona (lexema ligado+gramema ligado), pero se puede intercalar entre el lexema y el gramema de persona el de la vocal temática y los de tiempo y número:

am-o
am-a-mos

En el sustantivo tenemos los siguientes patrones más frecuentes:

1. Lexema libre: *ducha*.
2. Lexema ligado+gramema de género+gramema de número: *niñ-o-s*.
3. Lexema ligado+gramema derivativo+gramema de género+gramema de número: *niñ-it-o-s*.
4. Gramema preposicional+lexema ligado+ ...: *anti-american+ist+a+s*.

Construidas esas listas, tendremos finalmente los patrones morfológicos de las unidades palabra, propios de cada lengua. No hay, por lo tanto, *una* sola forma *palabra*, como tampoco hay una sola forma de sílaba o de morfema. La unidad palabra, existente en todas las lenguas, tiene formas diferentes en cada una de ellas.

6. LA UNIDAD PALABRA

Hemos podido reunir las tres condiciones necesarias para determinar la existencia de unidades *palabra* en cada lengua: sus características fonológicas (estructura silábica y existencia de funciones demarcativas de algunos fonemas o de algún rasgo suprasegmental), su característica semántica de unidad de denominación y sus características morfológicas, determinadas con ayuda de varios procedimientos. *Juntas, las tres condiciones se vuelven suficientes.*

Hemos agregado a esas tres condiciones, a la vez, una extensión de la concepción de la palabra, cuya procedencia es perceptual y económica, aun cuando la lingüística no se haya preocupado, hasta la fecha, por

someter a experimentación esas propiedades de la unidad palabra. Pero que no haya demostración experimental, no quiere decir que no exista; sobre todo, si lo que interesa a una lingüística verdaderamente dedicada a estudiar la realidad de las lenguas es considerar los fenómenos que estudia en toda su complejidad, en vez de descartarlos con ligeros juicios de sistematicidad.

Hemos visto también que en cada lengua hay varios tipos de unidad palabra, y hemos de estar dispuestos a encontrar otros, pero la tesis de este *Curso* es que, dados los criterios y procedimientos explicados, se pueden delimitar e identificar las palabras de cualquier lengua.

7. LA FORMA DEL CONTENIDO: MORFOLOGÍA Y SINTAXIS

Una conclusión más, que se puede sacar de este capítulo es que, ahora sí, es posible distinguir la morfología de la sintaxis: el morfema corresponde al cuarto nivel de análisis lingüístico, en donde se opera su segmentación y las pruebas de cohesión, en el que se descubren los diferentes esquemas morfotácticos y el orden de los morfemas que constituyen la palabra. La delimitación de la unidad palabra, que es el resultado obtenido en ese nivel, da lugar a un *quinto nivel de análisis lingüístico*. Corresponde a la morfología el estudio de los morfemas que conforman la palabra. La lexicología constituye el quinto nivel de análisis lingüístico. Después de ella, empieza la sintaxis.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

teoría de la doble articulación; elementos distintivos y significativos; base de la comparación; morfema; lexema; gramema; morfema libre y ligado; unidades de inventario abierto y cerrado; paradigma; morfotáctica; sufijo; prefijo; proclítico; enclítico; segmentación; cohesión; núcleo morfemático; prueba de intercalabilidad; capacidad de sustitución paradigmática; forma supletiva; paradigma polimórfico; composición; fraseologismo; orden de morfemas; sincategorema; categorema; palabra gramatical; palabra relacional.

OBRAS CONSULTADAS

ANDERSON, Stephen R., "Morphological Theory", en Frederick J. Newmeyer (ed.), *Linguistics. The Cambridge Survey, Vol. I: Linguistic Theory: Foundations*, Cambridge, Cambridge University Press, 1988, pp. 146-191.

ARONOFF, Mark, *Word Formation in Generative Grammar*, Linguistic Inquiry Monography 1, Massachussets, MIT Press, 1976.

BENVENISTE, Émile, *Ensayos de lingüística general*, México, Siglo XXI Editores, 1976.

BLOOMFIELD, Leonard, *Language*, Nueva York, Holt, Rinehart and Winston, 1933.

GREENBERG, Joshua, "The Word as a Linguistic Unit", en Charles Osgood (ed.), *Psycholinguistics. A Survey of Theory and Research*, Baltimore, Greenwood Press, 1954, pp. 66-71.

HEGER, Klaus, *Monem, Wort, Satz und Text*, Tubinga, Max Niemeyer Verlag, 1971.

HOCKETT, Charles P., *Curso de lingüística moderna*, traducción de E. Gregores y J.A. Suárez, Buenos Aires, Editorial de la Universidad de Buenos Aires, 1971.

JAKOBSON, Roman, y Morris HALLE, *Fundamentos del lenguaje*, Madrid, Ciencia Nueva, 1967.

LYONS, John, *Introducción en la lingüística Teórica*, traducción de Ramón Cerdá, Barcelona, Teide, 1977.

MARTINET, André, reseña a *Le mot*, de A. Rosetti, en *Word* 5,1 (1949), pp. 88-89.

MARTINET, André (dir.), *La lingüística. Guía alfabética*, Barcelona, Anagrama, 1975.

MATTHEWS, Peter H., *Morphology*, Cambridge, Cambridge University Press, 1974.

REY, Alain, *Le lexique: images et modèles : du dictionnaire à la lexicologie*, París, A. Colin, 1977.

4
EL SIGNIFICADO DE LA PALABRA

INTRODUCCIÓN

Al finalizar el capítulo anterior pudimos enunciar las tres condiciones necesarias, que juntas se vuelven suficientes, para delimitar lo que es una palabra en cualquier lengua. Incluso extendimos la definición de la unidad palabra a aquellos morfemas ligados, pertenecientes a inventarios cerrados, con poca o ninguna cohesión con la unidad de denominación, que restan en el quinto nivel de análisis una vez deslindadas las palabras de los tipos I a III, para poder darle su lugar a la capacidad que tienen los hablantes de una lengua de delimitar esos morfemas por separado e incluirlos como palabras en el nivel de descripción que les corresponde (que será el sexto o aun mayor, y cuyo estudio corresponde a la sintaxis). Con eso, terminamos de estudiar la forma del contenido de la palabra.

En este capítulo nos ocuparemos de la sustancia de contenido de la palabra, es decir, de su *significado* estrictamente hablando, para lo cual hemos de recuperar las explicaciones dadas en el segundo capítulo, en particular las del § 2.1 acerca de las bases psicológicas, cognoscitivas, de la unidad de denominación.

1. LA CONSTRUCCIÓN DEL SIGNIFICADO

Toda explicación de la manera en que se forma el significado de la palabra tiene que responder dos preguntas centrales. La primera es qué papel desempeña la aprehensión humana de la realidad circundante en la necesaria relación entre la palabra y aquello a lo que se refiere de la realidad (el *sentido*, según Hjelmslev); la segunda, cómo se relaciona esa aprehensión, que es ante todo individual, pues depende primeramente de las características físicas del cuerpo humano —los sentidos— con la

lengua histórica que transmite la sociedad a cada miembro de la comunidad lingüística.

En efecto, como habíamos visto en el capítulo 2, cada ser humano elabora esquemas de conocimiento a partir de sus acciones y sus experiencias. Esos esquemas obtienen su información de los sentidos: del tacto, del olfato, del gusto, del oído (gracias al cual el bebé comienza a relacionar su audición con su fonación, y a recibir estímulos constantes del habla de quienes lo rodean) y también de la vista. El esquema de conocimiento se construye, por eso, a partir de la información sensorial y la relevancia que van adquiriendo las experiencias sensoriales de las cosas, para llegar a reconocerlas como distintas unas de otras y nombrarlas. Si esto no fuera así, no podríamos explicarnos cómo es posible que nos entendamos al hablar de las cosas, y que entendamos que estamos hablando de "las mismas cosas" y no de otras.

1.1. La configuración perceptual

La biología nos enseña que el oído humano tiene capacidad para percibir ondas sonoras de cierto rango, dentro del cual se encuentran las del habla humana y multitud de ruidos y sonidos de la naturaleza. En cada especie animal, el rango de audición varía. Así por ejemplo, los perros pueden captar ondas ultrasónicas de una frecuencia que supera el rango audible para los humanos (son "ultrasónicas" para los humanos). Los silbatos utilizados para llamar la atención de los perros emiten esa clase de ondas. Lo mismo sucede con algunos aparatos modernos, que sirven para ahuyentar ratones e incluso insectos.

En cuanto a la vista, también el ojo humano puede captar cierta gama de colores, correspondiente al espectro de la luz blanca, que los arcoiris descomponen; más allá del violeta, en un extremo del espectro cromático del arcoiris, el ojo ya no puede percibir el ultravioleta; más allá del rojo, en el otro extremo, ya no puede percibir el infrarrojo. (Delimitaciones semejantes se pueden hacer para cada uno de los demás sentidos, aunque ya no sean de radiación, sino de transmisión de moléculas de unos cuerpos a otros, en el caso del olfato y del gusto.) Así que la percepción humana queda delimitada por las capacidades propias de los sentidos que intervienen en ella. Un objeto de la naturaleza puede tener muchas características vegetales, animales o minerales y emitir muchos tipos de radiación o de moléculas, pero lo que el ser humano percibe de él es sólo lo que queda dentro de las gamas de percepción de los sentidos.

La inteligencia humana tiene además la propiedad de tratar de asimilar a su organismo toda la información que recibe del mundo exterior o toda la experiencia sensible de su interior, como dolores o movimientos de los órganos. Cada vez que encuentra un estímulo que no había asimilado previamente, modifica sus esquemas de asimilación para adaptarse a él. Así, la relación entre el individuo y su experiencia sensible es de asimilación y de adaptación.

Esos procesos pueden verse desde dos perspectivas: la de las acciones humanas que involucran datos perceptuales del mundo sensible, y la de los resultados percibidos de esas acciones. Como se había dicho en el capítulo 2, las acciones que lleva a cabo el individuo desde sus primeras semanas de vida producen esquemas senso-motores de ellas en su memoria, que son los que vuelve a aplicar cuando repite cierta acción, y son los que se adaptan a las novedades que le ofrece una nueva acción. Por ejemplo, un bebé que manipula una sonaja, al principio, lo hará sin notar que se trata de una sonaja y que puede detener su propia mano para que deje de sonar o la puede agitar a su gusto. En el momento en que la sonaja y su sonido se vuelven notables, el esquema senso-motor de manejo de ella ya está establecido, y puede pasar a observar la propia sonaja y su sonido como elementos aislables de su esquema senso-motor previo. O imaginemos una experiencia que suele uno tener de niño: encuentra una piedra en el camino, la recoge y se da cuenta de su peso, su dureza y su informidad; en seguida, la lanza con la mano en alguna dirección. De esa acción llega uno a notar que, cada vez que lanza la piedra hacia adelante o hacia arriba, ésta describe una parábola hasta que cae al suelo; asimilada esta experiencia, pasa uno a calcular la fuerza que requiere lanzar una piedra de cierto peso y la altura a que conviene lanzarla para que llegue a cierto lugar. De allí seguimos al juego, pero también a la física de los proyectiles. Podemos decir que del esquema senso-motor producido por el manejo de la piedra, pasamos a un esquema más abstracto de la curva que describe la piedra al lanzarse y caer. Experiencias de esta clase son universales para los seres humanos, es decir, todo ser humano llega a tener esas experiencias, independientemente de la sociedad en que viva y de las características de su sociedad y su educación.

El psicólogo suizo Jean Piaget explicaba así la formación de la inteligencia: a base de esquemas senso-motores cada vez más complejos, de los cuales se llega a deslindar, en un momento dado, el conocimiento de las cosas, de las acciones y del resto de las experiencias que uno tiene a lo largo de la vida.

También podemos añadir que las propias características del cuerpo humano dan lugar a esquemas de conocimiento: la posición erecta del cuerpo y el hecho de que los ojos estén situados en un solo lado de la cabeza, correspondiente a la dirección en que se camina, dan lugar al esquema delante/detrás; el hecho de que la gravedad determine que la tierra quede a los pies y el cielo por encima de la cabeza, da lugar al esquema arriba/abajo. No así el esquema derecho/izquierdo o izquierdo/derecho que, como observan los maestros de escuela, no tiene el mismo carácter evidente de los dos anteriores, sino que requiere un entrenamiento social para construirlo.

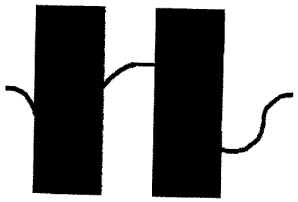
En el primer cuarto del siglo XX un grupo de psicólogos experimentales alemanes se dedicó a investigar cómo procede el conocimiento humano para adaptar los estímulos que reciben sus sentidos, en especial el de la vista (ya que tiene la ventaja de que resulta más fácil de estudiar en un laboratorio). Entre los muchos descubrimientos que llevó a cabo ese grupo, conviene tomar en cuenta dos: en primer lugar, que la inteligencia humana tiende a descubrir relaciones internas entre los objetos percibidos, que dan lugar a la formación de esquemas, de figuras de ellos, de *configuraciones*, que abstraen de diversas maneras las particularidades de los objetos. Como ese grupo de investigadores se dedicó a analizar la manera en que se reconocen las formas, bautizó su corriente científica como "psicología de la forma" o, con el vocablo alemán con que se lo conoce mejor, como "psicología de la Gestalt" (*Gestalt*, en alemán, quiere decir estructura y forma; por cierto, con ese antiguo morfema colectivizador *ge-* que encontramos en el capítulo anterior, en la etimología de *Gesellschaft*). A diferencia de la postura de Piaget, interesada por los procesos constructivos del conocimiento, por la *energeia* —¡la misma idea expuesta en el capítulo 2, § 1.2.!—, los psicólogos gestaltistas se interesaron por sus resultados —el *ergon*— orientados al conocimiento de los objetos mismos. Pero es posible y conveniente reunir los resultados de ambas escuelas de la investigación psicológica, para explicar los fenómenos verbales que nos interesan.

El otro descubrimiento que nos interesa aquí es que tales formas tienden a producirse privilegiando totalidades sobre elementos; es decir, que los seres humanos percibimos ante todo formas completas y no los elementos de que están formadas (a éstos llegamos después, mediante análisis).

Consideremos algunos ejemplos: si oímos una melodía, somos capaces de reconocerla aunque se cambie la tonalidad en que se toque, que altera las notas de que está compuesta, aunque no las proporciones que se dan entre ellas. Si no se reconociera la totalidad, cada alteración de una sola nota produciría una forma distinta para nuestra percepción. Gracias a esa capacidad, Juan Sebastián Bach pudo escribir la *Ofrenda musical* (BWV1079), una de sus obras maestras, compuesta por 16 variaciones de tono, organizadas en fugas, cánones, sonatas, etc., y ejecutables con muchos diferentes instrumentos, a partir de un solo tema, una sola melodía, dado al compositor por el emperador Federico el Grande de Prusia como un reto a su capacidad creadora.

Pero ilustrar esas investigaciones con datos musicales es más complicado en un texto como este. Por eso acudiremos a ejemplos de la visión (aunque, para la lingüística, es más importante la audición). Consideremos las ilustraciones que siguen, atendiendo a qué es aquello que nuestros ojos perciben en ellas:

En ésta, nos damos cuenta de que hay tres curvas separadas entre sí, sin que encontremos, espontáneamente, alguna relación entre ellas. Veamos, en cambio la siguiente ilustración:



En ésta, inmediatamente percibimos una curva continua que no logramos ver completa porque nos "estorban" dos barras negras. Hemos advertido inmediatamente la configuración de una curva sinuosa, a la que "reconocemos" una unidad, una continuidad, una forma. Lo que percibimos es una configuración, una *gestalt*. Todos los seres humanos tienen la misma capacidad para notar esa clase de configuraciones, independientemente de su educación. Incluso ha habido investigaciones que demuestran que muchos otros animales, como los pulpos (¡que son moluscos, por lo que ocupan un lugar comparativamente muy primitivo de la escala zoológica!), perciben *gestalten* (el plural de *gestalt* es *gestalten*, en alemán).

No nos extenderemos aquí sobre las propiedades de las *gestalten*, obtenidas por los psicólogos de esa escuela en sus laboratorios. Lo que nos interesa es establecer un punto de partida de la relación entre el ser humano y el mundo sensible, que nos servirá para comprender cómo se forma el significado.

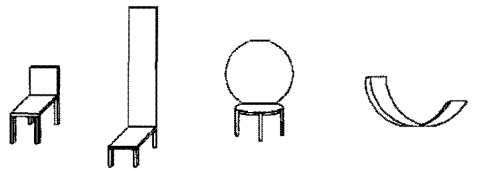
Otro conjunto de experimentos de la misma clase, desarrollado hacia la década de 1980 por la psicóloga estadounidense Eleanor Rosch tomó en consideración el campo del color: logró demostrar que, aun cuando el ojo humano percibe todo el continuo cromático del espectro de la luz —con las limitaciones antes señaladas—, ciertas longitudes de onda se convierten en los colores más claramente reconocidos por los seres humanos. Así por ejemplo, entre todos los matices del rojo, aquel que tiene mayor saturación, como el de la sangre arterial, se convierte en el "mejor representante" del color rojo. Así que la percepción del color, que es un fenómeno universal de los seres humanos, da lugar a ciertas preferencias que segmentan el continuo cromático en "focos" como el rojo, el amarillo, el azul, el verde, el violeta, etc. Y, para cada "color focal", se llega a elegir un "mejor representante". Según Rosch, estos colores focales son también fenómenos perceptuales universales a los seres humanos.

Podemos agregar a los ejemplos anteriores el de la *gestalt* del triángulo, cuyo mejor representante, según lo demostró Rosch, es el triángulo equilátero, y no el isósceles ni el escaleno. Eso se puede comprobar fácilmente ofreciéndole a una persona dibujos de las tres figuras geométricas y pidiéndole que elija la que le parezca "más triángulo" o un mejor ejemplo del triángulo. Como lo comprobó Rosch, incluso entre pueblos alejados del conocimiento de la geometría euclidiana, en Nueva Guinea, sus informantes seleccionaron el triángulo equilátero como mejor ejemplo de esa *gestalt*.

Las configuraciones que descubre la percepción en muchos objetos y la relación que hay entre ellas y sus "mejores representantes" es lo que Rosch decidió llamar *prototipo*.

Un prototipo, según Rosch, es una configuración sintética, no analizada, de atributos de un objeto, que tiene ciertas propiedades de proximidad, similitud, secuencia, agrupamiento, simetría, etc. entre los elementos que la conforman. Es un fenómeno perceptual que, como tal, no se puede reconocer en el ámbito de la lingüística, sino en el de la psicología.

Debemos tener cuidado de no confundir la *gestalt* percibida, el prototipo, que es un fenómeno psicológico, no directamente manifiesto a la inspección psicológica o lingüística, con su *mejor ejemplo*, una confusión que han sufrido muchos lingüistas cognoscitivistas contemporáneos. Esa confusión se debe a falta de claridad acerca del papel que tiene la ilustración del prototipo en un texto descriptivo o pedagógico. Por ejemplo, para estudiar la probable existencia de un prototipo de la silla (en las culturas de lengua española y muy probablemente en todas las que comparten la cultura occidental; aunque, según Rosch, si presentamos muchas sillas a personas de pueblos que no las tengan, llegarán a crear la misma *gestalt* y a definir el mismo prototipo que nosotros), ofrezco las ilustraciones siguientes:



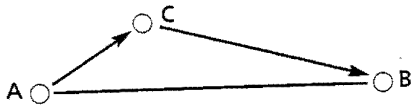
No debo confundir la ilustración que con mayor facilidad y espontaneidad reconozcan mis informantes como "la silla más típica" (que será la primera de la serie), con el prototipo mismo; sino que debo tener claro que se trata de su *mejor ejemplo*.

En relación con la multitud de objetos del mundo sensible y de las acciones o movimientos, tanto del mundo natural como del ser humano, la inteligencia humana crea esquemas de acción; cuando se trata de esquemas deslindados de ellas para reconocer objetos del mundo sensible, tenemos la clase de *gestalten* o prototipos que ha estudiado la psicología experimental; pero podemos hipotetizar que los esquemas de conocimiento de las acciones mismas también dan lugar a la formación de configuraciones perceptuales de la misma clase o semejantes a las *gestalten* o prototipos. Todos esos esquemas de conocimiento de acciones, que se forman en la inteligencia y se pueden considerar por separado de las maneras en que los nombramos, constituyen el *primer estrato de la formación del significado*.

El *primer estrato de la formación del significado* está constituido por configuraciones de carácter perceptual, construidas a partir de diversos esquemas de conocimiento.

En cuanto fenómenos de la percepción, los esquemas de conocimiento son universales y a ellos debemos, en buena parte, la capacidad de significar el mundo sensible con cualquier lengua, y para traducir esas experiencias de una lengua a otra.

Así por ejemplo, el movimiento o paso de un objeto C de la cercanía de un punto A a uno B da lugar a un esquema de acción, que podemos ilustrar como sigue:



(de nuevo: no hay que confundir el dibujo, que es una ilustración hecha para facilitar la comprensión, con el esquema de acción):

En seguida explicaremos cómo el español llega a significarlo con el verbo *dar*, pero también con *traer*, *llevar*, *ir* o incluso *entrar*; el inglés con *give*, *bring*, *go*, *come*; el alemán con *geben*, *bringen*, *gehen*, *kommen*, *eintreten*, etc., o el tzeltal con *yich'el*, *lok'an*, *xochon* o *xlok'on*.

1.2. El estereotipo

Con lo importantes que han sido los descubrimientos hechos por la psicología de la *gestalt* para comprender el funcionamiento de la inteligencia humana, y con lo importante para la lingüística que es encontrar el punto de unión entre los fenómenos que estudia con los que estudia la psicología, ese primer estrato del significado no pasa de ser una especie de esqueleto o de base perceptual del significado de algunas palabras, que conocemos y que manejamos. En los ejemplos recién dados, vemos que sí, el esquema dibujado puede representar varios verbos del español y de otras lenguas, pero no sirve para especificar sus diferencias, pues no significan lo mismo *dar* que *llevar*, *entrar* e *ir*, por un lado, o *recibir*, *traer*, *salir* y *venir*, por el otro (y correspondientemente, los verbos del inglés, del alemán y del tzeltal). Estos verbos podrán reducirse al mismo esquema de acciones, pero no son sinónimos de ninguna manera.

Como hemos venido insistiendo a lo largo de este libro, el ser humano aprende a hablar su lengua materna en plena inmersión en una gran cantidad de acciones, cada vez más complejas. Parte de esa complejidad estriba en que las acciones que lo llevan a distinguir las *gestalten*, los prototipos y los demás esquemas de acción pueden ir acompañadas por signos lingüísticos que las nombran y que contribuyen a construirlas. En la medida en que los esquemas de acción y de percepción se forman en multitud de situaciones verbales, los signos que los nombran no pueden ya separarse de su conocimiento. Tal inseparabilidad del signo y el esquema deja huellas profundas en su significado.

Analicemos entonces los verbos arriba enunciados: si estamos dentro de una habitación y hay una puerta que la separa de su exterior, la acción de moverse una persona del interior hacia el exterior, en español, se significa con *salir*; la acción contraria, con *entrar*, independientemente del lugar en que se encuentre la persona que la presencie. Así, decimos *entrar a la casa*, *salir de la casa*, ya sea que lo veamos desde dentro o desde fuera de la casa, porque el punto de vista significado por esos verbos del español fija la relación entre interior y exterior de la casa, que no depende ni del observador ni de quien habla. Si hemos de mover alguna cosa (C) de un punto A a un punto B, en español decimos: *llevar*, por ejemplo, *un pastel a la fiesta de unos amigos*; el punto A es nuestro punto de partida y el punto B es nuestro punto de llegada. Si quienes

hablan son nuestros amigos, en el punto B, dirán algo como "gracias por *traer* este pastel". En este caso, el punto de vista con que se enfoca la acción es el de quien se mueve en la dirección de B o el de quien aprecia el movimiento en B desde A. Por último, si hacemos pasar el objeto C de la posesión de A a la posesión de B, decimos en español *dar*. Aquí no interesan ni las posiciones de A y B, ni quien lo hace pasar y quien lo recibe: "te *doy* un regalo", "me *das* un regalo". "*Recibir* un regalo" subraya el punto de vista de quien lo recibe, exclusivamente.

En inglés, las acciones de entrada y salida se significan con *get in* y *get out*. Como en español, no tiene ningún papel el lugar en que se encuentra la persona que las presencia o que las realiza. La morfología de estos verbos es diferente: hay un elemento invariable *get*, seguido por dos preposiciones, *in* y *out*, que entran en cohesión débil con el lexema para formar las palabras *get in* y *get out* (no nos dejemos desorientar por la ortografía, de la que hablaremos en el capítulo 5). Son las preposiciones las que significan el 'hacia adentro' y el 'hacia afuera' de la casa. En alemán los verbos son semejantes a los del inglés: *eintreten* 'entrar' y *austreten* (+ *aus*, seguido del complemento) 'salir de alguna parte', con dos prefijos preposicionales que significan la dirección, aunque también se puede decir:

hineintreten 'entrar de A a B, visto desde A',
hereintreten 'entrar de A a B, visto desde B',
hinaustreten 'salir de A a B, visto desde A' y
heraustreten 'salir de A a B, visto desde B'

En inglés y en alemán esas diferencias con el español se vuelven a hacer manifiestas en la significación de las acciones de llevar y traer. En inglés *bring* y en alemán *bringen* quieren decir ambas cosas, y en alemán vuelven a actuar los gramemas preposicionales para darnos otros matices, por ejemplo, *einbringen* 'aportar algo a una discusión o a una tarea común', *mitbrigen* 'llevar algo a alguien' y otros más. En cuanto a la acción de dar, inglés y alemán se comportan de manera muy semejante al español: *give*, *geben*, respectivamente; de nuevo, en alemán, un prefijo preposicional distingue *eingeben* 'entregar o depositar algo a alguien o en algo' de *ausgeben* 'gastar o sacar dinero'.

Por último, el verbo ir en español significa que una persona cambie de un lugar A a otro B, ya sea a pie, ya sea en algún medio de transporte, considerado desde el punto A; con *venir* significamos el mismo movimiento, pero considerado desde el punto B de llegada. Así, yo *voy* a visitar a un amigo, y a él le alegra que *haya* yo *venido* a visitarlo. En inglés y alemán, el verbo que se utiliza para ese movimiento es *go* y *gehen* (en alemán, este verbo supone que lo hace uno a pie), pero si considero el movimiento entre A y B, y no solamente el que parte de A, entonces inglés y alemán optan por los verbos *come* y *kommen*, priorizando el punto de vista del lugar de llegada: *I will come to visit you* lo enuncio desde A a un amigo en B; ese amigo pudo haberme dicho *When will you come to visit me?* ('¿cuándo vendrás a visitarme?'). En alemán es semejante, e incluso se dice *Ich komme mit dir* para significar 'voy contigo de A a B, cuando nuestro interlocutor está en B'.

El tzeltal distingue la posición del observador, de manera semejante al alemán: *yich'el talel* quiere decir 'traer algo al punto B', en tanto que *yich'el bael* significa 'llevar algo de A a B' (el verbo *yich'el* quiere decir, sin los morfemas que definen la posición de salida o de llegada, 'tener algo'). También como el alemán, moverse desde un punto A hacia un punto B, independientemente de que A o B se encuentren fuera o dentro de una casa, se dice *lok'an tal*, mientras que el mismo movimiento, significado desde B, se dice *ochan tal*: por ejemplo, *ya me xochon tal* 'voy a entrar a donde tú estás', frente a *ya me xlok'on tal* 'voy a salir a donde tú estás'.[1]

Como podemos ver en estos ejemplos, la configuración de esas acciones es la misma, y gracias a ella podemos comparar los verbos con que se significan en las cuatro lenguas; de otra manera, no podríamos explicarnos por qué somos capaces de reunir esos verbos en la comparación y tampoco seríamos capaces de traducir de una lengua a la otra. Esa es la importancia de la base perceptual que nos ofrece la configuración de las acciones consideradas.

Pero también nos damos cuenta de que los significados de esos verbos "dicen más" de las acciones que lo que percibimos con nuestros sentidos, reducidos a su papel orgánico. Las direcciones en que se dan las acciones, los puntos de vista de quien las presencia, quien las ejecuta o quien recibe lo aportado no forman parte de los correspondientes esquemas, sino que son diferencias significativas que tienen las lenguas. ¿Cómo llegaron estas cuatro lenguas a encarnar esas diferencias? Espe-

[1] Agradezco los ejemplos del tzeltal al maestro Roberto Santiz Gómez.

rarfamos que nos lo explicara un estudio etimológico de esos verbos, aunque en la mayor parte de los casos, o quizá en todos, su origen quede en la oscuridad completa del pasado humano y sólo podamos calar un poco en él.

Hemos venido tomando como una de las bases epistemológicas de este *Curso* la lengua histórica, es decir, la lengua real, no las abstracciones que se suele hacer en lingüística para descubrir en ellas sólo su sistema o deducir sus meras virtualidades. Apoyarnos en ese concepto de la lengua histórica equivale a no negar, sino a tomar en cuenta el hecho de que las lenguas se transmiten socialmente; que llegan a cada individuo ya conformadas, como resultado de larguísimas tradiciones verbales. Son esas tradiciones, construidas a lo largo de los siglos y de manera inconsciente para cada miembro de una sociedad, las que han ido creando las diferencias significativas que encontramos en el significado de las palabras ejemplificadas.

Pues lo que hace a una lengua histórica es una memoria transmitida, de generación en generación, de aquellos productos verbales que han resultado valiosos para hacer inteligibles las experiencias de los miembros de una sociedad a todos sus integrantes. Quizá el estudio etimológico no encuentre datos suficientes como para poder mostrar las maneras en que se constituyeron los significados de las parejas de verbos que hemos venido analizando, pero lo cierto es que esos significados no corresponden al mero esquema de conocimiento, a la mera forma de una *gestalt*.

Aquel elemento del significado de una palabra que se forma en una lengua histórica concreta, como resultado de una larga valoración de sus experiencias verbales, y que vuelve inteligible la palabra para los miembros de la comunidad lingüística es lo que, siguiendo al filósofo inglés Hilary Putnam, llamaremos *estereotipo*. Podemos considerarlo el *segundo estrato de la formación del significado*.

Para aclarar este concepto, tomemos algunos ejemplos más, sobre todo uno de los que aduce el propio Putnam para explicar su concepto de estereotipo: si le preguntamos a un físico qué es la *electricidad*, podemos comprobar, ante todo, que el físico es perfectamente capaz de entender la palabra *electricidad*, por lo que comprende muy bien nuestra pregunta. Pero muy probablemente nos responda que, bien a bien, no hay una explicación única de lo que es la electricidad, sino que varias teorías físicas son capaces de explicarla de distinta manera, a partir de diferentes elementos.

Si eso es así, quiere decir que el significado de la palabra *electricidad* no nos ha sido dado por la física a los hablantes de español (y de todas las lenguas en que existe una palabra correspondiente), es decir, no proviene de una definición física de lo que es ese fenómeno, sino de una experiencia previamente socializada de lo que es la electricidad. Por eso el físico entiende su significado como nosotros y sabe, como nosotros, que la electricidad circula por cables y alambres, que se produce en instalaciones termoeléctricas, hidroeléctricas, eólicas e incluso nucleoeléctricas, y que es una poderosa energía que produce movimiento en los motores, luz en los focos, calor en las estufas, etc.; sabe que da toques y que, en ciertos casos, puede matar a un ser viviente. Incluso, el físico puede dar instrucciones a un ingeniero acerca de cómo construir un motor que la genere o un instrumento que la transforme. ¿Cómo se formó este significado de la palabra *electricidad*? A base de las experiencias de las sociedades que la usan: fue la cultura griega la que plasmó su observación de la electricidad estática que se producía al frotar con un paño de lana una piedra de ámbar (el origen de la palabra es *elektron*, que significa 'ámbar'), pero no comenzó a investigarse ese fenómeno sino en el siglo XVI.

A mediados del XVIII se inventaron aislantes y conductores para utilizarla; Benjamín Franklin descubrió en esa época que el rayo era eléctrico, y Faraday inventó en la misma época el motor eléctrico y el primer transformador; sin embargo, sólo en 1864 James Maxwell ofreció las ecuaciones con que se pudo teorizar, finalmente, el fenómeno del electromagnetismo. Es decir, primero se formó un significado estereotípico de la palabra *electricidad* y siglos más tarde se pudo proponer una definición teórica del fenómeno. La experiencia de uso de la electricidad ha creado un significado de la palabra, independientemente de la clase de fenómeno físico-químico de que se trate.

En español decimos que *el Sol sale* y *el Sol se pone*; sale por el este, se pone por el oeste. Esta manera de hablar, recibida por nosotros desde hace muchos siglos, corresponde a una experiencia común de todo ser

humano: ver aparecer por el horizonte todas las mañanas el Sol, verlo *nacer*, seguir su recorrido por el cielo y verlo ponerse o *morir* todas las tardes por el oeste. Según esa experiencia, tan fácilmente metaforizable con la experiencia de la vida humana, *el Sol se mueve*. El hecho de que usemos esas palabras normalmente para darnos a entender, a pesar de que, desde Kepler, sepamos que el Sol está fijo, para efecto de todo cálculo importante desde la Tierra, indica que nuestra tradición verbal ha formado un *estereotipo* de la relación entre la Tierra y el Sol, que es el que guía nuestro entendimiento. Ese estereotipo, heredado desde hace siglos, es falso en relación con la teoría astronómica, pero es verdadero desde el punto de vista de nuestra capacidad para darnos a entender verbalmente.

Hablando del mismo Sol, sabemos que una parte del significado de la palabra *sol* es que se trata de un astro que brilla de día. Parte del significado de la palabra *estrella*, en cambio, es que se trata de un astro que brilla de noche. De ahí que la oposición entre *sol* y *estrella* sea casi *antonímica*. Esos significados son estereotípicos. No obstante, la astronomía moderna nos enseña que el Sol es una estrella, pues todas las estrellas son cuerpos celestes caracterizados por un estado de combustión permanente de los elementos químicos que las forman; una teoría a la que ha llegado la astronomía estudiando, precisamente, el Sol y midiendo después la energía procedente de las estrellas.

Nuestros estereotipos del Sol y las estrellas organizan nuestra expresión, incluso la de los astrónomos que los estudian, aunque desde el punto de vista científico, sean falsos. Quizá pronto los niños de escuela, educados con conocimiento científico, dejen de oponer el Sol y las estrellas, y los dos significados que manejamos caigan en desuso y terminen por desaparecer. En ese momento, los estereotipos cambian o desaparecen. *Un estereotipo, por lo tanto, es temporalmente verdadero*, tanto como la lengua histórica lo siga manteniendo.

Pero no sólo eso: los estereotipos, por ser los que hacen comprensible nuestra comunicación, determinan la *corrección de nuestras expresiones*. Si alguien dice "Anoche vi millones de soles brillando en el cielo", la reacción normal de sus interlocutores es corregirlo, y decirle que lo que vio brillando fueron estrellas; si uno ha aprendido que el Sol es realmente una estrella, sonreirá, al menos, por el exceso de precisión de quien lo ha dicho.

Los animales más comunes para cada sociedad dan lugar generalmente a la formación de estereotipos. Un gato, por ejemplo, difícilmente se puede distinguir sólo por su forma, pues comparte gran parte de ella con los perros, los tigres y muchos otros cuadrúpedos. Pero la experiencia social con el gato da lugar a un significado de la palabra *gato* en donde no cuenta precisamente que sea un felino, sino un animal doméstico, de ojos muy brillantes en la oscuridad, de andar elástico y silencioso, de piel aterciopelada, que maúlla, que caza ratones, que es enemigo del perro y tiene siete vidas. Ese es su estereotipo, y gracias a él, que constituye el significado de la palabra *gato*, podemos entender expresiones como "tener más vidas que un gato" ('sobrevivir muchas situaciones peligrosas'), "andar como perros y gatos" ('pelear o discutir continuamente dos personas'), "ponerle el cascabel al gato" ('atreverse a enfrentar a una persona más poderosa que uno mismo'), etc.

El estereotipo tiene, en consecuencia, tres características: *a*] se construye socialmente y es el que determina la pertinencia social del significado; *b*] es temporalmente verdadero, y *c*] determina la corrección de muchas expresiones.

1.3. El papel de la cultura en la formación del significado

Ya dijimos antes que muchas palabras tienen una base perceptual, correspondiente al primer estrato de formación del significado; *ir* y *venir*, en español, corresponden a la misma configuración de acciones, pero las diferencias importantes en sus significados se producen en otro plano, correspondiente a la experiencia histórica de la comunidad lingüística, que pone en juego diversas relaciones de enunciación de la acción, de dirección, de situación en relación con algún punto de referencia, etc. De la electricidad, del Sol o del gato debe haber un esquema de conocimiento, que probablemente no sea del orden de la *gestalt*; cualquier ser humano puede distinguir esos tres objetos, pero lo que constituye el significado de las palabras *electricidad*, *sol* y *gato* no es del orden de la percepción, sino de la transmisión histórica de un conocimiento verbal, en que se han

Hay varias maneras de comprender lo que es cultura y de definir su significado. Hoy en día muchos autores la entienden como el conjunto de las costumbres o de los hábitos de una comunidad determinada. Este es un concepto de cultura que podríamos clasificar como "descriptivo". Con ese concepto se habla, por ejemplo de "la cultura de los niños de la calle" o de "la cultura de los narcotraficantes". Otro concepto, que es el que manejamos en este libro, supone que la cultura conlleva una valoración de experiencias, lo cual la hace muy diferente de la mera colección de hábitos. Cuando se habla de "la cultura tolteca" se entiende el conjunto de los mitos y las artes que nos heredó a los mexicanos, lo cual supone una valorización de esos mitos y artes. En consecuencia, la pobre vida de los niños de la calle no llega a crear una cultura, y tampoco las costumbres violentas, corruptas e inmorales de los narcotraficantes.

destacado las características más importantes de esos objetos para una comunidad lingüística particular. En estos tres casos, el estrato del estereotipo al que corresponden, determina más el significado.

Una lengua histórica es también una acumulación y selección de experiencias lingüísticas, que avanza valorando lo ya dicho y seleccionando lo que vale la pena seguir diciendo, en relación con nuevas experiencias que han de adquirir pertinencia y sentido para todos los miembros de la comunidad. Este proceso es lo que solemos entender por *cultura*.

Volvamos a nuestro ejemplo del verbo *dar* en español. Hay un conjunto de verbos más, que pueden corresponder al mismo esquema de acción: *donar*, *regalar*, *prestar*, *conceder*, *otorgar*, *heredar* y unos pocos más. Sus significados no son solamente del orden de la percepción, sino también del del estereotipo, pero además, han venido formándose en la significación de acciones valiosas para la cultura de la lengua española y a lo largo de los siglos; generalmente sobre la base de distinciones ya existentes en latín, pero siempre haciendo resaltar ciertos rasgos nuevos, como se puede ver si uno estudia su desarrollo histórico.

Comencemos por tomar en cuenta el significado de *donar*: la acción de hacer pasar algo (C) de A a B es la misma que la de *dar*, pero la diferencia estriba en que lo que se da son objetos que constituyen un bien para quien los recibe; *se dan dones*. No se *donan* problemas, deudas o males, sino bienes, dinero, virtudes. Aquí, la cultura de la lengua española tomó una voz latina para crear una diferencia entre el simple *dar*, que es neutro, básico, y el *donar*, que hace hincapié en los dones.

En *regalar*, su significado es dar algo a alguien para que haga uso de él libremente y como un acto de generosidad de quien da, algo muy distinto del *prestar*, que significa dar algo a alguien temporalmente, con el compromiso de devolverlo. En el significado de *regalar* cuenta más la

generosidad y la libertad de quien da los objetos, aunque éstos también han de considerarse bienes apetecibles.

Ahora comparemos *otorgar* con los verbos anteriores: se *otorga* algo de carácter más elevado, más abstracto, más espiritual, considerado de mucho valor; pero además, se hace como muestra de verdadera gracia de quien lo otorga, de magnanimidad, pues no está obligado a hacerlo, ni quien recibe lo otorgado puede alegar ningún derecho a recibirlo. Así por ejemplo, un médico no puede escribir al comité del Premio Nobel para pedir que *se lo otorguen*; el comité puede hacerlo como un premio gratuito y desinteresado a un médico, para reconocer el valor de su trabajo. Entre *regalar* y *otorgar* hay una sutil diferencia, que corresponde al papel social de quien regala u otorga, y al carácter más concreto —como dinero, una joya— de lo regalado, frente al más elevado de lo otorgado, que generalmente se subraya con una ceremonia especial: "te *regalo* un coche", pero "el Senado te *otorga* una condecoración".

El significado de *conceder*, al contrario del de *otorgar*, supone siempre una solicitud previa para que alguien dé algo a alguien: se *concede* un préstamo cuando uno lo solicitó previamente. (Hay periodistas o burócratas que, a causa del mayor respeto que se atribuye a quien otorga algo, piensan que, cuando un *servidor público concede* una entrevista, la otorga; pequeña sutileza del adocenamiento.)

Por último, *heredar* significa dar algo de su propiedad a otra persona, generalmente emparentada con uno, como regalo antes o después de morir. En *heredar*, lo que cuenta es el parentesco o cercanía con quien da la herencia, y su muerte ineludible. Todas estas diferencias de significado, que forman parte de la cultura verbal de la lengua española nacieron para hacer distinciones valiosas o importantes entre diversas acciones consistentes en hacer pasar C de A a B. Muchas de las distinciones se gestaron en el ámbito jurídico, en donde era necesario hacerlas para poder juzgarlas y legalizarlas. La mera percepción no basta para aprender los significados de estos verbos, sino que sólo se transmiten a lo largo de una práctica verbal educada, que nunca termina para cada ser humano.

Los significados que se crean en la cultura de una lengua constituyen el tercer estrato de su formación.

No es por eso extraño, aunque sí doloroso, que personas que nacen privadas del oído, tengan tantas dificultades para aprender los significados de todas estas palabras. Aquéllas cuyo estrato de significado más determinante es el de las configuraciones perceptuales, se aprenden fácilmente; las otras en que se forman estereotipos y, sobre todo, las que deben su significado a la cultura verbal, requieren, en primer lugar, una larga inmersión en las acciones verbales de la comunidad; en segundo, un proceso educativo, que es el principal objetivo de la educación. Una persona que goza del sentido del oído tiene la capacidad más necesaria para aprender el significado de las palabras; quien no lo tiene, requiere una educación especial, a base de una lengua de señas y del aprendizaje de la lectura y la escritura.

Conviene ampliar la ejemplificación del tercer estrato de formación del significado, para poder apreciar cómo éste no sólo puede no tener una base perceptual, sino que también depende, centralmente, de la cultura verbal. Tomemos por caso la palabra *dragón*. Como los dragones no existen en la naturaleza, nadie podrá verlos o tener contacto con ellos. En griego δράκων y en latín *draco* significaban 'serpiente', aparentemente derivado de una raíz indoeuropea que quería decir 'brillo, esplendor'. Nacido el animal y la palabra en leyendas antiguas, por mucho tiempo se creyó en su existencia en la naturaleza.

Aunque durante la Edad Media muchos caballeros tomaban su figura para aterrorizar a sus enemigos, y lo representaban, como ahora, con alas de murciélago, fauces de las que salía fuego, garras y cola terminada en una flecha, el *Diccionario de Autoridades*, que es el primero que compuso la Real Academia de la Lengua en España, entre 1713 y 1726, definía su significado como sigue: "Serpiente de muchos años, que con el tiempo crece y tiene un cuerpo grande y grueso. De ordinario le figuran con pies y alas para distinguirle de la serpiente y algunos autores dicen que con la edad las cría". El diccionario no podía distinguir, en aquella época, entre la existencia o no del dragón en la naturaleza, por lo que, cautamente, atribuía a otros la descripción de sus características.

Hoy en día la sociedad sabe que no existen en la naturaleza, sin embargo, ha formado un estereotipo, a base de esas mismas leyendas y las representaciones que ha heredado de ellos *en la cultura*. Cualquier niño que pueda oír cuentos en que aparezcan dragones o que pueda ver ilustraciones de ellos, nos dirá que un dragón es una especie de gran serpiente, con

alas de murciélago, que echa fuego por la boca. Incluso, a partir de ese cúmulo de experiencias de origen exclusivamente verbal, luego ayudadas por las ilustraciones de los cuentos, llega a formar un prototipo del dragón, pero lo importante es recalcar su procedencia exclusiva de la cultura.

Podemos ahora resumir lo que es la construcción del significado de una palabra: sobre la base de esquemas de acción y de conocimiento, la inteligencia humana lo mismo llega a crear configuraciones perceptuales de objetos del mundo sensible, que a elaborar características y rasgos de experiencias que resultan pertinentes y valiosas para la comunidad lingüística; en esta construcción, el significado que se transmite verbalmente (en acciones verbales o por escrito) se sedimenta como producto de la cultura, orientado a hacer cada vez más inteligibles todos los acontecimientos y experiencias de la vida humana. Si esto se puede decir del español o del francés, o del alemán, se puede afirmar igualmente del náhuatl, de las lenguas zapotecas o del tzeltal.

Si el ejemplo de la palabra *dragón* puede parecer pueril (o sea, de niños), lo mismo que afirmamos de ella lo podemos decir de muchas otras palabras que nos parecen "muy serias", como *democracia, soberanía, libertad, caridad* o *amistad*. Ninguna de ellas tiene en su significado un estrato perceptual y ni siquiera estereotípico, sino que se ha construido exclusivamente en el ámbito de la cultura. Por eso hay tantas discusiones acerca de lo que debe entenderse por esas palabras, pues no hay unos objetos, igualmente reconocibles por cualquier ser humano, o unas ideas flotando en el aire, que sean "la democracia", "la soberanía" o "la libertad" a los que simplemente "representen" esas palabras. (De ahí la necedad de varios gobernantes estadounidenses, que creen posible exigir a pueblos muy diferentes, de otras culturas, que vivan, de la noche a la mañana, en una "democracia" como la suya, o que puedan imponérsela a base de bombas.) Los significados de esas palabras se han venido gestando mediante complejas elaboraciones intelectuales, en las cuales los grandes filósofos de la historia de Occidente han tenido un papel fundamental; son significados construidos por la actividad verbal, que se aprenden como parte de una educación en la comunidad lingüística.

1.4. El significado especializado

Conviene agregar un elemento más, cada vez más importante para las lenguas conforme crece la dependencia de nuestra civilización de la invención científica y los instrumentos de la técnica. La lengua histórica, por las características que hemos venido estudiando en ella, ofrece los medios para construir de manera comprensible para los demás cualquier pensamiento o cualquier experiencia, por compleja y singular

que sea. Estamos acostumbrados, por las características de nuestra cultura de la lengua, a entender el papel que tiene la lengua en la literatura. Es un lugar común afirmar que el *Quijote* es una de las más grandes obras de la literatura en lengua española, al grado de que parece muy difícil que sus traducciones a otras lenguas logren reproducir la experiencia verbal que nos ofrece a los hispanohablantes. Eso mismo se piensa de la traducción de poesía: el traductor puede *recrear* el poema, pero no vertirlo exactamente de una lengua a la otra, sin perder elementos esenciales de su belleza. Que la literatura sea una construcción verbal, nadie lo pone en duda.

En cambio, se piensa que el conocimiento técnico y científico no tiene nada que ver con la lengua en que se hace inteligible; incluso ha habido quienes creen que las palabras que tienen un significado especializado, como *hidrógeno*, *bit* o *fulereno*, no forman parte de la lengua, sino que toman de ella su significante exclusivamente (lo que, como vimos desde el primer capítulo de este libro, es imposible, si aceptamos la definición del signo lingüístico enseñada por Saussure). La realidad es otra: que esas palabras no sólo tienen formas (de la expresión y del contenido) correspondientes a las características de cada lengua, sino que se originan, la mayor parte de las veces, en la cultura de una lengua y aprovechan los significados que tienen para construir con ellos sus significados especializados.

Así por ejemplo *hidrógeno*: la existencia de este gas se conocía por lo menos desde finales del siglo XVII, aunque no se supiera cuál era su naturaleza. Un autor alemán, Stahl, lo bautizó en 1702 como *flogisto*, tomando la palabra griega φλογιστοξ, que quiere decir 'inflamable' (no es una palabra científica), porque lo que observaban estos primeros químicos de la historia era que se podía quemar. Antoine Lavoisier, ya en el siglo XIX, que siguió haciendo observaciones en su laboratorio notó que, como resultado de su calentamiento, aparecía agua, por lo que, volviendo al patrimonio verbal griego, optó por llamarlo *hidrógeno*, es decir, 'hacedor o engendrador de agua' (*hidros* significa 'agua' y *genos* 'hacedor, engendrador').

La palabra resultó entonces una creación nueva, culta, formada en el conocimiento de la lengua griega, que entró a formar parte de muchas lenguas apenas en el siglo XIX. Como se ve, este término técnico de la química tuvo su origen en una cultura verbal. Lo mismo podemos decir

de *bit*, una palabra mucho más moderna, que nació cuando se construyó la computadora electrónica, que trabaja a base de dígitos binarios; es decir, a base de representaciones de acciones de conexión y desconexión de un circuito. En inglés, *binary digit* dio lugar al acrónimo *bit*, pero aprovechando que en esa lengua *bit* significa 'pedacito', sus inventores hicieron un juego de sentido muy útil para que la palabra fuera rápidamente comprendida por su comunidad lingüística.

Por último, los *fulerenos*, que son moléculas muy complejas de la química orgánica, recibieron su nombre del apellido del arquitecto angloamericano Buckminster-Fuller, porque el químico que los creó se dio cuenta de que su estructura es geodésica, como las que solía construir ese arquitecto. No fue estrictamente la cultura verbal inglesa el origen de la palabra, pero sí la cultura angloamericana.

Así que no hay distinción alguna en la manera en que se forma el significado de las palabras comunes y la manera en que se forma el significado de los términos técnicos, pero sí la hay en el proceso de significación que siguen unas y otros. En las comunes, el significado se forma en la actividad verbal de la comunidad, atendiendo a elementos que resulten valiosos y pertinentes para la cultura, pero sin una dirección precisada por nadie; en los términos técnicos hay una voluntad expresa por crear palabras o por crear significados que *designen*, lo más unívocamente posible, algún objeto al que hay que referirse (véase el capítulo 9).

La estipulación o la definición precisa de un significado de una palabra, orientadas a una designación unívoca de un objeto, constituyen el *cuarto estrato de la formación del significado*: *el estrato del término técnico*.

En todas las lenguas se forman los significados de la misma manera; a veces predomina el estrato de las configuraciones perceptuales, a veces se le suma el del estereotipo y a veces incluso el de la cultura verbal y el del interés especializado; a veces no se forman estereotipos, sino que los significados se construyen en una cuidadosa red discursiva, que va delimitando los significados de las palabras, como sucede con *democracia* o *amistad*.

2. LA POLISEMIA

Hemos venido hablando del significado como sustancia del contenido de las palabras, para mostrar la manera en que se forma entre la experiencia sensorial individual y la actividad verbal de la sociedad. Falta ahora tratar un fenómeno que es evidente para cualquier persona que se detenga a reflexionar un poco acerca de sus propias palabras. Este fenómeno consiste en que *cada palabra puede tener más de un significado*. Volvamos a la palabra *gato*: si su significado principal es el de 'animal doméstico, cuadrúpedo, de piel aterciopelada, ojos brillantes, andar silencioso, que maúlla, que es enemigo de los ratones y los perros, etc.', en el español mexicano entendemos también por *gato* un instrumento mecánico o hidráulico que sirve para levantar objetos muy pesados, en especial, los coches cuando hay que componerles una rueda; llamamos *gato*, además, al juego que consiste en llenar antes que el contrincante tres casillas contiguas de una red de nueve, en dirección horizontal, vertical o diagonal, y también significamos con *gato*, peyorativamente, a un sirviente. Es decir que la palabra *gato* tiene cuatro significados, que se pueden delimitar con completa precisión.

La *polisemia* es el fenómeno que consiste en que una palabra tenga, cuando se la considera en aislamiento, es decir, fuera de cualquier contexto, más de un significado (lo cual se muestra claramente en los diccionarios).

La polisemia es un fenómeno de la naturaleza significativa de la palabra; a partir de la acción verbal inmediata de los individuos se configura en la evolución semántica de la lengua. Es muy importante tomar en cuenta esta característica de la polisemia, que es inherente a la manera de ser de las palabras de cualquier lengua, porque suelen producirse dos confusiones acerca de ella: por un lado, al agrupar el término técnico de *polisemia* con los de *sinonimia*, *homonimia* y *antonimia*; por el otro, al confundirlo con la *ambigüedad* y la *vaguedad*.

Como se ha mostrado en el parágrafo anterior, la polisemia de una palabra se produce como resultado de la capacidad que ofrece cualquier lengua para significar cualquier nueva experiencia, valiosa para

la comunidad lingüística. Es un *fenómeno de la palabra* en sí. Por el contrario, la *sinonimia*, que definiremos provisionalmente como la igualdad de significado de dos o más palabras; la *homonimia*, que consiste en la igualdad de los significantes de dos o más palabras con distinto significado, y la *antonimia*, que se suele entender como una oposición simétrica entre los significados de dos o más palabras, son *fenómenos del léxico*, es decir, del conjunto de las palabras de la lengua histórica (una definición precisa de estos tres fenómenos se verá en el capítulo 9).

Algunos lingüistas confunden el fenómeno de la *polisemia* con la *vaguedad* y la *ambigüedad*. Es necesario aclarar que se trata de fenómenos diferentes. La polisemia es un resultado del uso histórico de cada palabra que, como hemos visto, se presta a significar cualquier experiencia, por nueva que sea para cada hablante. Por ejemplo: cuando llegó Cristóbal Colón a las islas de las Antillas por primera vez, o sea, cuando los europeos descubrieron que había otra tierra al otro lado del océano, suponía que había alcanzado las costas de India, de la que habían tenido muchas noticias en la Europa de su época. Al ver a los habitantes originarios de esas islas y comprobar su existencia en el resto de nuestro continente, los europeos los llamaron *indios*.

La palabra utilizada era el antropónimo de los naturales de India; tiempo más tarde se reconoció que el continente americano era diferente del asiático, pero ya se había generalizado el uso de la palabra *indio*, que por eso adquirió un nuevo significado; se volvió polisémica. Hoy en día, *indio* es el gentilicio de los habitantes de India (*hindú* es el nombre de los indios de religión brahmánica, no de todos los habitantes de India) y el antropónimo de los habitantes originarios de América. No es que la palabra *indio* sea ambigua o vaga, pues podemos precisar sus dos significados.

La ambigüedad y la vaguedad son defectos de la proposición y del texto, no de la palabra (ni del diccionario, a menos que sea de mala calidad). Uno puede siempre utilizar la palabra en la oración con ambigüedad, si no la conoce bien, e incluso hacerlo así para producir algún efecto con su discurso. Si digo, en América, "me encontré con un indio en la calle", el contexto cultural me hará suponer que se trata de una persona que desciende de los pueblos originarios de América, pero mi proposición es ambigua, pues podría tratarse de un ciudadano de India.

Este juego con la ambigüedad de las proposiciones es uno de los mecanismos centrales del chiste. Aquel letrero en un camión que decía: "te cambio mi llanta nueva por tu vieja", juega con dos significados diferentes y precisos de la palabra *vieja*: 'que ha vivido muchos años, que ha servido por muchos años y resiente el paso del tiempo', lo cual se puede predicar tanto de las llantas como de las mujeres, y el que se ha añadido a *vieja* para querer decir 'esposa'.

La *vaguedad*, por su parte, consiste en la falta de precisión de un discurso o un texto determinado para alcanzar a formular claramente el sentido de lo que se dice. No daremos ejemplos de ella, pues es fácil de reconocer en diversos textos o discursos, y muchas veces es un medio que utiliza una persona para no afirmar algo o para no comprometerse con algo.

2.1. La metáfora y la polisemia

La polisemia de una palabra es resultado de un proceso de significación por el cual una persona aprovecha alguna de las características de expresión o de contenido de una palabra para hablar de una nueva experiencia que logra transmitir adecuadamente al resto de los miembros de su comunidad y que, por eso mismo, se vuelve pertinente para la memoria social de las experiencias compartidas.

La retórica antigua, que se concebía como un catálogo de procedimientos de significación aplicables en diferentes tipos de discurso y para diferentes necesidades discursivas socialmente definidas y aceptadas, sitúa la metáfora entre los llamados *tropos* o "figuras de pensamiento", junto con la *metonimia* y otras "figuras" como la *sinécdoque*, la *catacresis*, etc. Como el objetivo de este capítulo y de este libro no es la retórica, bastará con dejar señalado el papel de estas "figuras de pensamiento" en la formación de los significados de las palabras y dar a la metáfora un reconocimiento privilegiado, como "figura de pensamiento" que engloba a todas las demás, y sobre todo, como el fenómeno por excelencia de la creación de nuevos significados de las palabras.

Por ejemplo, en el español nacional de la península ibérica, a las jóvenes que se contrata por unas horas para cuidar a los hijos de poca edad de una familia las llaman *canguros* (en México, son *niñeras*). No cuesta trabajo entender el proceso de significación que añadió ese nuevo significado a la palabra *canguro*: si el estereotipo de los marsupiales australianos que conocemos como canguros consiste en el hecho de que cargan a sus crías en una bolsa de su cuerpo, es fácil entender que a una

joven, encargada por unas horas de los hijos pequeños de una familia, "le apliquen" el mismo nombre del animal.

En el español coloquial de México, *canguro* tiene también un segundo significado, para nombrar una especie de mochila, que se lleva delante del cuerpo, para cargar en ella a un niño pequeño. La retórica antigua llamaba a estos procesos de significación *metáforas*. Es fácil imaginarse que la metáfora es el mecanismo más importante de la significación en una lengua histórica y que a esa propiedad semántica de las lenguas se deba la flexibilidad y la capacidad que tiene toda lengua para "trabajar lo inexpresable hasta que se vuelva expresable", como lo entendió Louis Hjelmslev.

3. CONCLUSIÓN

Con este capítulo terminamos la caracterización completa de la unidad palabra y los métodos con los que podemos definirla en cada lengua particular. Puesto que todavía en nuestra época la mayor parte de las lenguas del mundo son habladas y no tienen sistemas de escritura, la caracterización recién terminada nos permite definir la unidad *palabra* como fenómeno oral. En el siguiente capítulo nos ocuparemos de la palabra escrita.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

configuración perceptual; *gestalt*; prototipo; estereotipo; mejor representante del prototipo; estratos de formación del significado; cultura verbal; término técnico; polisemia; metáfora; tropo; figura de pensamiento.

OBRAS CONSULTADAS

KOFFKA, Kurt, *Principios de psicología de la forma*, Buenos Aires, Paidós, 1973.
LARA, Luis Fernando, *Teoría del diccionario monolingüe*, México, El Colegio de México, 1997.
——, *Ensayos de teoría semántica. Lengua natural y lenguajes científicos*, México, El Colegio de México, 2001.
PIAGET, Jean, *La formación de la inteligencia en el niño*, Madrid, Aguilar, 1969.
PUTNAM, Hilary, *Mind, Language and Reality. Philosophical papers II*, Cambridge, Cambridge University Press, 1975.
ROSCH, Eleanor, "Human Categorization", en N. Warren (ed.), *Studies in Cross-cultural Psychology*, Londres, Academic Press, 1977, t. i, pp. 1-49.

5
LA PALABRA ESCRITA

INTRODUCCIÓN

En los cuatro capítulos anteriores exploramos todas las características pertinentes para el análisis lingüístico de las unidades *palabra*. Como resultado de los tres primeros, enunciamos las condiciones necesarias y suficientes para determinar su existencia en cualquier lengua. Es especialmente importante resaltar que la unidad *palabra* es un fenómeno universal de las lenguas, es decir, se puede demostrar empíricamente la existencia de la *palabra* en cualquier lengua, aunque sus características formales en cada una de ellas sean diferentes; es decir, la unidad *palabra* de cada lengua es contingente a las estructuras propias de cada lengua. No podría ser de otra manera. Lo mismo se puede decir de las unidades fonema, sílaba y morfema. Por eso, se distinguen consonantes de vocales, diferentes patrones silábicos, diferentes paradigmas de morfemas y diferentes esquemas oracionales.

En el capítulo 4 nos ocupamos de la sustancia del contenido de la unidad *palabra* y mostramos de qué manera se relaciona la palabra con la realidad circundante y con las experiencias a las que los seres humanos tratamos de dar sentido. Aprovechamos la enseñanza que hasta ahora nos ha podido ofrecer la psicología experimental, en cuanto se refiere al reconocimiento perceptual de objetos y acciones en el mundo sensible; la situamos como parte del proceso adaptativo humano con que Jean Piaget explica la formación de la inteligencia, y deslindamos el carácter fundamentalmente verbal —es decir, procedente de acciones verbales de significación— del estereotipo y de la transmisión cultural en la formación del significado. Por último, tratamos dos fenómenos característicos del significado de la palabra: la polisemia y el papel que tiene en su formación la metáfora, que agrega significados a los ya existentes en la lengua histórica.

Ahora pasaremos a considerar el papel de la escritura en la determinación de la unidad palabra. Al referirnos a la escritura, ya no se puede hablar de ella como si fuera otra característica universal de la palabra, pues son menos las lenguas del mundo que se escriben y muchas más aquellas que siguen siendo solamente orales. Sin embargo, para las lenguas que tienen escritura, ésta llega a ser su última determinación; es decir, lo que reconocemos como palabras en español, en alemán, en inglés, en tzeltal, en náhuatl moderno, etc., tiene un cuño insoslayablemente escrito. A la vez, toda escritura que desee elaborarse para una lengua que no la haya desarrollado, tendrá varias de las características que a continuación explicaremos.

1. CONCEPTOS BÁSICOS

La escritura ha sido inveteradamente soslayada por la lingüística moderna. En su afán por acercarse a la diversidad de las lenguas del mundo de la manera más directa, eludiendo al máximo la influencia de la tradición gramatical normativa de las principales lenguas de Occidente, la lingüística moderna concibió la escritura como un elemento de esa tradición que contribuía a deformar la observación de la oralidad; y además, asumió que la escritura no era sino una mera trascripción de la forma de la expresión de una lengua, considerada como una cadena de fonemas, en el caso de la escritura alfabética, o un simbolismo convencional del *sentido* (no de la sustancia del contenido, pues ésta es siempre inherente a cada lengua), en el caso de las escrituras ideográficas, como el chino, o jeroglíficas, como el maya, sin ningún interés para la lingüística. Por ese motivo, durante decenas de años, la lingüística se desentendió de la escritura; sólo en la última década, varios lingüistas han comenzado a interesarse por la escritura como fenómeno complejo y a llamar la atención sobre sus características, tanto en su realidad contemporánea como en los procesos de aprendizaje de la lectura y la escritura de los seres humanos.

Debido a eso es necesario establecer previamente algunos conceptos básicos que ayuden a comprender adecuadamente las características principales de la escritura, para poder pasar después a analizar su papel en la determinación de la unidad *palabra* en lenguas que, como el español, la han incorporado centralmente a su propia configuración.

1.1. Sistema de escritura

Un sistema de escritura puede ser *ideográfico* o *glotográfico*. Es decir, puede estar construido para transmitir unidades de sentido, independientemente de las características de la lengua de quienes lo utilicen (y, por lo tanto, sus símbolos gráficos son completamente independientes de cualquier lengua), o puede construirse para transmitir unidades de una lengua particular, ya sean palabras, morfemas, sílabas o fonemas (y entonces, sus trazos representan alguna característica formal de la lengua en cuestión).

Gloto- es un morfema tomado del griego γλοσσα, que quiere decir 'lengua'.

El ejemplo por antonomasia de una escritura ideográfica es el chino; en la ilustración que sigue, los cuatro caracteres mostrados simbolizan 'agua', 'pájaro', 'hombre' y 'sol'; para poder leer un texto chino no hace falta saber hablar la lengua, pues se puede aprender la escritura sin la lengua de la cultura en que se creó el sistema (pero sí estudiando esa cultura). Prueba de ello es que el japonés, una lengua de tronco diferente al del chino, adoptó el mismo sistema para su cultura, los llamados "kanji", que suele combinar con una escritura silábica propia, las escrituras "hiragana" y "katakana".

Muchos tratadistas de la escritura china consideran que estos caracteres tienen un origen *icónico* (*icónico*, también del griego εἰχών, quiere decir 'retrato fiel de alguna cosa') que hoy día, y sin estar educados en la cultura china, nos resulta imposible reconocer (quizá pudiéramos imaginar las plumas en la parte inferior del carácter de 'pájaro' y el pico al lado derecho suyo; en el de 'hombre', quizá tratáramos de interpretar las piernas unidas al tronco, y en el de 'sol' ya nos resultaría imposible imaginar un primitivo círculo con rayos).

Los trazos de los ideogramas chinos se han vuelto abstracciones de dibujos que originariamente "retrataban" objetos de la realidad (la verdad es que el sistema de escritura china es una compleja combinación de ideogramas y símbolos, algunos de carácter glotográfico, pero como

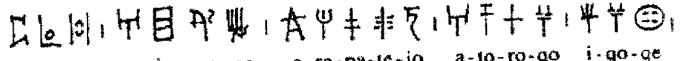
sol hombre pájaro agua

aquí se trata sólo de ilustrar lo que son los sistemas ideográficos, basta con ello).

Los sistemas *glotográficos*, en cambio, como decíamos antes, representan signos de una lengua particular (palabras, morfemas) o elementos de segunda articulación, como sílabas y cadenas de fonemas. En estos casos, es necesario conocer la lengua a que pertenece el sistema para poder comprender el significado de lo escrito. Cuando el sistema representa unidades de segunda articulación, se trata de un *sistema fonográfico*. La escritura del español forma parte de esta clase de sistemas.

Cabe observar que, como lo demuestran varios estudiosos de la historia de las escrituras, en la genealogía de los sistemas glotográficos se encuentra la *unidad de denominación* (véase el capítulo 2) como base para su creación y que aun en los ideográficos, puesto que las "ideas" no se pueden reconocer independientemente de su comunicabilidad, hay un proceso de denominación de la realidad en el origen de sus símbolos, como se puede colegir con sólo considerar los cuatro ideogramas chinos mostrados antes.

Tomemos por caso, de los sistemas glotográficos, el llamado "lineal B", utilizado por la civilización micénica del sur de Grecia entre los siglos XVI y XIII a.C., correspondiente a una forma antigua de la lengua griega, y tomado de una civilización anterior, de lengua desconocida, llamada minoica, cuyos restos de escritura se denominan "lineal A" y siguen siendo ininterpretables. Pues bien, en la escritura micénica se separaban *unidades de denominación* mediante pequeños trazos verticales entre ellas, con objeto de que la escritura fuera inteligible.

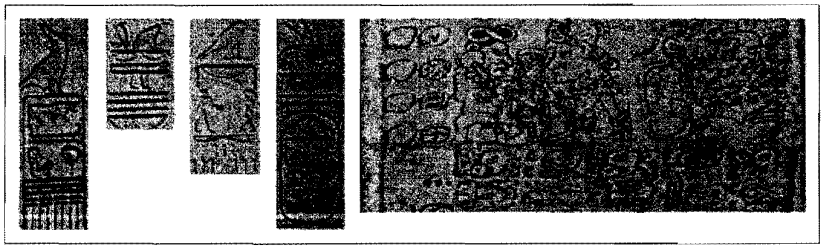


ta-ra-nu	a-ja-me-no	e-re-pa-te-jo	a-to-ro-qo	i-qo-qe
$t^hrānus$	aiaimenos	elep^{h}anteiois	ant^{h}rōk^{w}ōi	hik^{w}k^{w}ōik^{w}e

La ilustración proviene del libro de Sampson citado en la bibliografía, p. 104.

(La primera línea en escritura latina indica que la lineal B era un sistema silábico; en la segunda, está la versión que reconstruye las palabras griegas, cuyo significado era: "un escabel con ilustraciones de un caballo, un hombre y un pulpo".)

Quizá lo mismo se pueda afirmar de las escrituras jeroglíficas, una variedad de las ideográficas, como el egipcio antiguo y el maya. No es extraño que sea así, pues la escritura es un medio para referir a la realidad, que nació *representándola*, y la mejor manera de representarla es aislar un objeto y ostentar su "retrato" a la vista de los demás, ya sea mediante íconos o por medio de símbolos abstraídos de ellos, o también, mediante sus *nombres* correspondientes; así pues, podemos suponer que las unidades *palabra* de cada lengua tuvieron algo que ver en la creación de todos los sistemas de escritura.



1.2. Las características del sistema de escritura fonográfico

Pero lo que nos interesa en este capítulo es analizar el sistema de escritura que utilizan el español y muchas lenguas europeas, así como muchas de las amerindias actuales. Quien se interese por otros sistemas de escritura, como los recién aludidos, podrá encontrar análisis e información en las obras que se listan al final de este capítulo.

Un sistema de escritura como el del español está formado por: *a*] un conjunto de trazos visibles en una superficie de contraste, socialmente instituidos; *b*] un conjunto de reglas de representación de formas lingüísticas, y *c*] un conjunto de reglas de complexión del propio sistema. A continuación pasaremos a explicar cada uno de estos elementos.

1.2.1. Los trazos

Poca gente se da cuenta de que los trazos de nuestra escritura, las *letras*, no se inventaron desde la nada para utilizarlas en la escritura del español, del francés, del alemán o incluso del latín y del griego, sino que forman parte de una herencia milenaria, que ha venido transmitiéndose desde los siglos X u VIII a.C., a partir del momento en que la cultura griega adaptó la escritura fenicia a su propia lengua. La lengua fenicia era una lengua semítica, como lo son el árabe y el hebreo. Una particularidad de esas lenguas es que las vocales varían en el interior de la palabra de acuerdo con su contexto morfológico y sintáctico, a diferencia de las lenguas indoeuropeas, en que las vocales son relativamente fijas en la palabra, aunque haya varios procesos morfo-fonológicos que actúan sobre ellas (por ejemplo, en inglés, *goose* 'ganso' y *geese* 'gansos'; en alemán, *rad* 'rueda' y *räder* 'ruedas').

Dada esa variación, la representación de las vocales no se hacía completamente necesaria, por lo que la escritura fenicia sólo tenía trazos para las consonantes (las escrituras árabe y hebrea tampoco tuvieron trazos de representación de las vocales por mucho tiempo, y no los tienen actualmente en cierta clase de textos).

Cuando los griegos tomaron la escritura fenicia para su propia lengua, adoptaron todas las consonantes fenicias, incluso en el mismo "orden alfabético" fenicio, pero se vieron en la necesidad de adaptar —no crear— algunas letras para representar sus vocales. Así por ejemplo, la letra <N> (utilizaremos los ángulos para indicar que se trata de letras y no de fonemas), llamada *nu* por los griegos, fue primitivamente una representación icónica de una serpiente (nombrada *nohas* en lengua fenicia), de la cual provino la forma que dio lugar a la letra fenicia, llamada *nun* (en hebreo, la letra נ), tomando su trazo serpenteante para representar el fonema /n/ con que comenzaba la palabra *nohas*. En cambio, la letra *yod* fenicia, proveniente de un ícono de la mano, nombrada *yad*, representaba un fonema semiconsonante /j/, y pasó a representar la vocal /i/ del griego, con lo que dio lugar a la letra *iota* griega <I>. De allí, a través del latín, vienen nuestras letras *ene* e i. Lo interesante de esa historia es que ni los griegos, ni los romanos, ni los pueblos europeos herederos de la cultura romana se plantearon inventar nuevas letras, sino que se concretaron a cambiar la representación de las letras fenicias para adaptarlas a sus propios fonemas y, en muy pocos casos, crearon letras nuevas, como la *eñe* del español.

Esta pequeña incursión por la historia de la escritura muestra, por un lado, que los trazos del sistema de escritura latino, heredado por todas las lenguas del Occidente europeo, evangelizadas desde Roma, así como los del cirílico (llamado así porque lo creó San Cirilo para lenguas como el ruso o el búlgaro) de las lenguas del Oriente europeo, evangelizadas desde Bizancio, tienen un determinante carácter tradicional, que los ha instituido entre todas las sociedades que los utilizan; por otro lado, que esas letras no nacieron como una convención abstracta y relativamente caprichosa para adecuarlas a los fonemas de las lenguas, sino que *se disociaron de los nombres* de los objetos que representaban icónicamente en sus orígenes, tomando el ícono del objeto representado como representación del primer fonema de la palabra con que se lo significaba, lo que viene a ser una buena muestra del papel central de la palabra incluso para la escritura fonográfica.

Los trazos en sí, que han venido variando mucho a lo largo de la historia, son objeto de mucho interés para la paleografía, dedicada a la lectura de documentos antiguos; para la historia de los instrumentos de escritura, como los cálamos latinos y las tablillas de cera, las plumas, los pergaminos y los códices, el papel, las tintas, los bolígrafos, las máquinas de escribir y, hoy en día, los pixeles de las pantallas de las computadoras; para el diseño de letras y tipos de imprenta, y para el diseño de letras de rasgos y tamaños adecuados para la enseñanza de la escritura.

Podemos llamar *letras* a las *gestalten* de los trazos que constituyen el sistema de escritura. Así por ejemplo, las *letras* A o F pueden trazarse de maneras muy diferentes, como A, A, F, F, etc., dando lugar a diferentes *tipos* de escritura.

En cambio, podemos llamar *grafía* a cada letra o conjunto de letras correspondientes a un fonema o grupo de fonemas. Así por ejemplo, al fonema nasal palatal /ñ/ del español le corresponde en nuestro sistema de escritura la grafía <ñ>, en tanto que a fonemas semejantes de otras lenguas corresponden otras grafías: <nh> en portugués, <ny> en catalán o <gn> en francés e italiano; al fonema palatal africado sordo /ch/ le corresponde la grafía <ch> en español, en tanto que a un fonema semejante del italiano, por ejemplo, le corresponde la grafía <ci>, o del vasco <tx>; la letra <x>, en cambio, es grafía del grupo de fonemas /ks/ en es-

pañol (por ejemplo, en *éxito*), mientras que en el español mexicano es grafía, además de /ks/, del sonido correspondiente a un fonema náhuatl /sh/ en *mixiote*, y del fonema /x/ en *México* y *Xalapa*.

1.2.2. Las reglas de representación

Dado el hecho histórico de que los sistemas de escritura del español y de las otras lenguas que hemos venido considerando preferentemente (por ser las que más nos interesan en nuestro horizonte cultural inmediato) se manifiestan, *en principio*, como instrumentos de codificación de los fonemas de los signos de estas lenguas, se pueden considerar *sistemas fonográficos*, aunque, como veremos más adelante, no lo son exclusivamente. El criterio básico de constitución de un sistema fonográfico es el del llamado *principio fonológico*.

1.2.2.1. El principio fonológico

El *principio fonológico* consiste en la necesidad de que a cada fonema de una lengua corresponda una grafía de su escritura, y a cada grafía le corresponda siempre el mismo fonema de la lengua. Es decir, la relación entre fonema y grafía debe ser *biunívoca*. Esta es la *primera regla* de representación del sistema fonográfico de escritura.

En el caso del español, en donde el acento de intensidad es distintivo —es decir, tiene valor fonológico—, se esperaría que, a pesar de que se trate de un rasgo suprasegmental, de características diferentes a las de los fonemas, ese acento formara parte igualmente del sistema y se indicara en todas las palabras. Después veremos por qué no es así.

1.2.2.2. Orden lineal y dirección

Es característico de los sistemas glotográficos, en general, que precisamente por el hecho de que representan, bien signos o bien unidades de segunda articulación de cada lengua particular, se escriban siguiendo un

orden lineal, en correspondencia con el orden de nuestro discurso, que es un hecho temporal: siempre que hablamos, un signo va después de otro. Ese orden de la escritura puede ser horizontal, de derecha a izquierda, como en árabe, o de izquierda a derecha, como en español; puede ser vertical, generalmente de arriba hacia abajo, y puede seguir una dirección hasta el final de la línea o de la columna, y dar la vuelta allí mismo, invirtiendo su dirección en la siguiente línea o columna, un orden llamado *bustrofedón*, lo cual quiere decir 'giro del buey', un orden comparable con la vuelta que da la yunta al terminar de arar cada surco.

La *segunda regla* del sistema de escritura del español, y de las demás lenguas que hemos venido tratando en este libro, es que el orden sea de izquierda a derecha, de arriba a abajo, y que se interrumpa para volver a comenzar en el margen izquierdo de cada línea.

Las dos reglas anteriores son *constitutivas* del sistema fonográfico. A base de ellas se puede construir un sistema de escritura fonográfica para cualquier lengua.

Muchos tratadistas de la escritura, desde Gonzalo Correas en el siglo XVII español hasta Raúl Ávila en México o Jesús Mosterín en España, contemporáneamente, se han esforzado por lograr que la gran comunidad hispanohablante modifique su escritura para que se atenga por completo a la primera regla.

Lo mismo ha sucedido en otras comunidades lingüísticas. La inglesa y la francesa, por ejemplo, que también tienen, en principio, un sistema fonográfico de escritura, han visto esfuerzos periódicos por ajustarlo a su lengua, para corregir una escritura que constantemente se aparta de la primera regla.

Esos esfuerzos no han fructificado, porque se basan en una racionalidad analítica que no corresponde a la realidad de la escritura, que es un objeto cultural y lingüísticamente tan complejo, pues hay que considerar en ella, en primer lugar, el hecho de que nuestros sistemas de escritura se conformaron como instituciones herederas de poderosas tradiciones escriturales, tan antiguas como la misma escritura griega y latina, y no son resultado de convenciones puntuales, arbitrarias y perfectamente racionales, correspondientes a un análisis lingüístico previo, basado en el conocimiento del sistema fonológico de cada lengua, a partir del cual se hubiera producido una verdadera codificación de sus fonemas.

En segundo lugar, el hecho de que su conformación ha evolucionado orientada por el papel social que han tenido la lectura y la escritura desde la época latina hasta nuestros días. La buena voluntad de esos tratadistas de la reforma de nuestro sistema de escritura se estrella con su compleja realidad.

1.2.3. Las reglas de complexión del sistema

1.2.3.1. El acento

Decíamos antes, en el § 1.2.2.1, que el acento, por ser distintivo en español, debería formar parte de las correspondencias que establece el principio fonológico entre fonemas y grafías, pero que no ha sido así. Como sabemos, en español distinguimos entre *acento prosódico* y *acento ortográfico*.

El acento prosódico es precisamente el acento que tiene carácter distintivo en español, pero bien sabemos que ese acento no se representa en todos los casos, pues palabras como *rojo* o *añil*, que tienen acento prosódico en sus penúltima y última sílabas, no se escriben con acento ortográfico, lo cual muestra una discrepancia entre éste y aquél. Las reglas de acentuación ortográfica alteran el principio fonológico con otras reglas, que modifican su aplicación mecánica rigurosa y definen en qué condiciones se escribe la tilde con que marcamos el acento. Las reglas más importantes son:

1. No se acentúan gráficamente los monosílabos como *de*, *con*, *el*, *don*, *sin*, etcétera.

2. Se escribe el acento solamente cuando una palabra aguda termine en vocal, ene o ese: *maná*, *miré*, *comí*, *sillón*, *camión*, *autobús*, *anís*.

3. Se escribe el acento en una palabra grave solamente si termina en consonante que no sea ene o ese (nótese que esta regla es recíproca de la anterior): *fácil*, *árbol*, *útil*, *cárter*, *córner*, *cáncer*, *huésped*, *lápiz*.

4. Se escribe el acento en toda palabra esdrújula o sobresdrújula: *depósito*, *brújula*, *característica*, *teológico*, *dictámenes*, *regímenes*.

5. Se escribe el acento para distinguir *homógrafos* —es decir, signos diferentes que se podrían escribir de la misma manera, de acuerdo con el principio fonológico, como *de* (preposición) y *dé* (imperativo de *dar*); *mas* (conjunción adversativa) y *más* (adjetivo o adverbio); *solo* (adjetivo) y *sólo* (adverbio); *el* (artículo) y *él* (pronombre), etc. (Trataremos esta quinta regla después, en el § 1.2.3.4.)

Las reglas de acentuación que dan lugar al acento ortográfico modifican la aplicación de la primera regla constitutiva del sistema fonográfico por motivos que ya no corresponden a la representación fonográfica, sino a una *economía interna del propio sistema*, puesto que sólo en los casos que preveen las cinco reglas ortográficas hay posibilidad de confusión *en la lectura*.

Las reglas de acentuación ortográfica vienen a ser *reglas reguladoras* del sistema de escritura; no son reglas constituyentes del sistema, como las dos primeras, vistas antes, sino que sólo regulan su aplicación, lo que en este caso obedece a un principio de economía interna del sistema, pues reducen de manera importante la necesidad de codificar el acento a aquellos casos en que la lectura podría dar lugar a confusiones: acentuar los monosílabos no hace falta, pues no pueden leerse de otra manera (cuando se producen homografías, como entre la conjunción *mas* y el adjetivo o adverbio *más*, entra en función la regla ortográfica); si lo más común en español es que sus palabras sean graves, es decir, acentuadas prosódicamente en la penúltima sílaba, las reglas del acento ortográfico aprovechan esa regularidad para volver innecesaria la codificación de su acento, y sólo definen las condiciones en que aquella regularidad se rompe, como en los casos previstos de las graves terminadas en consonante que no sea ene o ese, las agudas terminadas en vocal, ene o ese, y las esdrújulas.

Como vemos, esa economía del propio sistema tiene un carácter diferente al de las dos reglas que constituyen el sistema de escritura. Llamaremos a todas esas reglas de economía interna del sistema, *reglas de complexión*. Las reglas de acentuación ortográfica son *reglas de complexión* del sistema.

1.2.3.2. La configuración morfológica

Las escrituras francesa e inglesa son dos buenos ejemplos de cómo un sistema fonográfico se altera, lo cual ha dado lugar a muchas dificultades en sus respectivas comunidades lingüísticas e inspira muchos debates, dirigidos ya a modificar sus sistemas para ajustarlos rigurosamente al principio fonológico, ya a poner en duda si se trata realmente de sistemas fonográficos.

Nina Catach, una de las principales estudiosas de la escritura del francés, ha comprobado que el 80 u 85% de las grafías del sistema de escritura francés se usa para codificar efectivamente sus fonemas, mientras que el resto de sus usos corresponde a la codificación de otros aspectos importantes de la lengua; esa es una razón suficiente

para considerar que el sistema de escritura del francés es, a pesar de todo, un sistema fonográfico de base, aunque complementado en un 3 a 6%, por usos suplementarios de las letras que dan información sintagmática (de posición) y paradigmática (flexional o derivativa); por un 12 a 13% de grafías con valor diacrítico, así como de valor simbólico etimológico e histórico, y por un 3 a 6%, de lo que esta autora considera verdaderos ideogramas (o sea, semejantes a los de la escritura china) como *eau* 'agua', pronunciado [o] abierta, en que la forma de la palabra escrita es insegmentable, tanto en su expresión como en su contenido.

Tomemos por ejemplo precisamente el caso de los dos fonemas /o/ abierto y cerrado del francés. En el sistema de escritura ambos se codifican con la grafía <o>. En *robe* 'vestido' y *joli* 'bonito', la grafía codifica la vocal abierta; en *rose* 'rosa' y *pot* 'recipiente, tarro', la cerrada. Sin embargo, la misma distinción fonológica se puede codificar de otras maneras: en *saute* 'salto o cambio brusco' se trata de la vocal /ǫ/ abierta, mientras que en *sotte* 'tonta', de la cerrada. En *eau* tenemos un *trigrama* (una grafía compuesta por tres letras) para codificar la vocal abierta, que en este caso es, además, una palabra, como acabamos de decir. La misma letra <o>, con un acento circunflejo como diacrítico, codifica el mismo fonema, como en *côte* 'costa' o en la interjección exclamativa: "*ô ciel!*" '¡oh, cielo!'. Combinada con otras letras (forma *digramas*) codifica la vocal nasal /õ/: *on*; la vocal /u/ cerrada: *cou* 'cuello'; *ou*, conjunción disyuntiva (y *où*, adverbio de lugar, con acento grave como diacrítico); la semiconsonante /w/: *oui* 'sí', *pingouin* 'pingüino'. La correspondencia entre todas estas grafías y los fonemas /o/ es materia de regulación ortográfica.

No entraremos en explicaciones de carácter histórico para aclarar el sistema de escritura del francés y su ortografía, pues no es ese el objetivo de este capítulo ni de este libro; lo que hay que tomar en cuenta es la manera en que ha quedado conformada la escritura francesa.

Si analizamos *saute* y *sotte*, en que la distinción fonológica de vocal abierta y cerrada se ve codificada por el digrama <au> y por la regla ortográfica que señala que una vocal seguida por una consonante doble debe leerse como cerrada, podemos reconocer que en este caso el principio fonológico se cumple parcialmente, porque la codificación de ambas vocales no se hace con la misma grafía —como en *robe* y *rose*— sino con letras que corresponden a dos reglas diferentes:

a] /ǫ/ abierta → <au>
b] /ǫ/ cerrada → <o> seguida de consonante doble

Reglas semejantes podemos suponer en la codificación de la nasal /õ/, en que la escritura francesa optó por codificarla con *on*, en vez de ponerle una tilde, como en portugués, o una diéresis, como se hace actualmente en la /ã/ de *hñähñu*, el nombre que dan los otomíes a su lengua.

Si comparamos ahora *sot* 'tonto' /sǫ/ con *sotte* 'tonta' /sǫt:/, en que la vocal de la forma masculina es abierta y la de la forma femenina es cerrada, comprobamos la existencia de una regla morfo-fonológica del francés, que tiene su efecto en la complexión del sistema. Es decir, notamos el reflejo de un fenómeno más complejo que el que está en la base del principio fonológico. El sistema sigue siendo fonográfico, sólo que ahora predomina una distinción gráfica basada en la necesidad de mostrar la relación entre las dos formas sobre la base escrita de la femenina, que es *más informativa* que la masculina.

Otra escritura de /ǫ/ abierta es *aux*, que encontramos como forma que une la preposición *à* y el artículo masculino plural: *aux châteaux* 'a los castillos', y como forma del plural masculino de palabras terminadas en *-au*. En la lengua hablada, [ǫ shatǫ] con ambas vocales /o/ abiertas, se necesita el contexto para poder distinguir si se habla de un solo castillo: *au château* 'en el castillo, al castillo' o de varios. El sistema de escritura permite hacer la distinción e introduce una repetición de la forma plural para aclarar la interpretación. En este caso, el sistema adquiere *redundancia*, lo que es una condición de eficacia de todo sistema de información.

1.2.3.3. Informatividad y redundancia

La teoría de la información, que se desarrolló en el ámbito de la ingeniería de las comunicaciones, pero cuya influencia se ha vuelto aun más importante al relacionarse con la termodinámica y con la genética, nos ofrece dos conceptos muy valiosos para poder comprender sistemas complejos como los de las lenguas y los sistemas de escritura: *información* y *redundancia*. Tomemos un ejemplo sencillo para explicarlos:

Supongamos que en nuestra casa instalamos un dispositivo junto a la puerta de entrada, que se acciona cada vez que se abre la puerta, y se hace perceptible ya sea tocando un timbre o encendiendo una luz. Cada vez que se acciona, recibimos la información de que la puerta se acaba de abrir. Pero supongamos que un día falla la corriente eléctrica, por lo que se abre la puerta, pero ni suena el timbre ni se enciende la luz. No tenemos manera de enterarnos de lo que ha sucedido. El pequeño sistema del dispositivo es cien por ciento informativo... siempre que haya corriente eléctrica. Como nos es muy necesario saber cuándo se abre la puerta —porque tememos que se meta un ladrón en la casa, por ejemplo— le agregamos una luz (no un timbre, pues sería muy molesto) que está encendida todo el tiempo, para indicar que hay corriente eléctrica.

De esa forma, si la luz se apaga, sabemos que el dispositivo no funcionará al abrir la puerta y tomaremos nuestras precauciones. Al ampliar de esa manera el sistema, la luz que indica que hay corriente eléctrica, nos da una primera información, y ésta garantiza la segunda información —de que se abre la puerta—. El sistema tiene un grado de redundancia. Si no solamente instalamos una luz, sino además un timbre para garantizar la transmisión de la segunda información, aumentamos la redundancia del sistema y, aunque hay repetición, nos resulta más eficaz, desde el punto de vista de la información que recibimos.

Todas las lenguas tienen grados de redundancia para que sus expresiones sean informativas. En español, por ejemplo, la indicación del plural en una oración como "vinieron a visitarme todos mis amigos" es redundante, pues se manifiesta en el sufijo de número del primer verbo y en los del adjetivo, el pronombre y el sustantivo. En algunos dialectos del español en que se aspira la /s/ final y se pueden producir oraciones como "[todo ello comen papa]", la única manera de saber que se trata de un plural es gracias al sufijo de número en el verbo, por lo que no hay redundancia, a menos —como sucede en realidad— de que las vocales finales de las palabras se abran y sustituyan el sufijo de número por una oposición entre vocal cerrada y abierta: [todo ello] 'todo ello' y [papa] frente a [todǫ elle] 'todos ellos' y [papa].

En el sistema de escritura del francés formas como *-aux* sirven para aumentar la informatividad del texto escrito, en comparación con el discurso hablado, y para agregar una redundancia que, debido a la evolución histórica de la lengua francesa, se ha trasladado a otros elementos del contexto. En el del español, si tomamos en cuenta dialectos en que se aspira la ese final, sucede lo mismo, aunque se trate de una sola letra, correspondiente a un solo fonema. La fonografía se completa con una *morfografía*.

1.2.3.4. Los diacríticos

En el § 1.2.3.1 incluimos una quinta regla reguladora del sistema de escritura del español, por la cual se introduce el acento ortográfico para diferenciar funciones de palabras que se escriben de la misma manera, como el adjetivo *solo* y el adverbio *sólo*. A propósito de la escritura francesa, mostramos cómo con un acento circunflejo sobre la vocal <o> se distingue una interjección exclamativa <ô>. El acento circunflejo sirve para distinguir por escrito esta interjección, que al hablar es solamente una [o] quizá un poco alargada y pronunciada con un tono de voz más alto de lo normal. El fonema *-o* de palabras como *château*, *plateau* 'planicie', *oiseau* 'pájaro' se codifica con <au> para el singular y con <aux> para distinguir por escrito el plural: *châteaux*, *plateaux*, *oiseaux*. La escritura diferente de esos dos fonemas facilita su reconocimiento por escrito, en tanto que al hablar, el francohablante necesita otros elementos del contexto que, mediante redundancia, le permitan distinguir el singular del plural, como el artículo [le] con /ę/ cerrada, frente a [le] con /ę/ abierta: *le château* frente a *les châteaux*, *le plateau* frente a *les plateaux*, y la sonorización de /s/ además de la /e/ abierta para pronunciar [lęz wasó], *les oiseaux*. La introducción de letras y otros trazos como diacríticos en un sistema de escritura forma parte de su complexión; como vemos en la escritura francesa, esos diacríticos vuelven su sistema de escritura un instrumento muy eficaz de reconocimiento de la lengua, facilitando la precisión del mensaje escrito.

Algo semejante sucede con el inglés: según lo explica Sampson, la economía de la escritura inglesa requiere que todo morfema léxico se codifique al menos con tres letras: *see* 'ver', *bee* 'abeja', *egg* 'huevo', en tanto que los morfemas de inventario cerrado puede tener menos letras: *me* 'yo', *be* 'ser o estar'. Para distinguir los homófonos (palabras que se pronuncian de la misma manera, pero tienen significados distintos) como en /rait/, el sistema de escritura ofrece *right* 'correcto', *rite* 'rito', *write* 'escribir' y *wright* 'obrero'.

Forman parte de las reglas de complexión del sistema de escritura todas las letras que no codifican fonemas, sino distinciones de carácter morfológico, sintáctico y de significado.

1.2.3.5. La conservación etimológica

Un elemento importante de los medios con que cuenta un sistema de escritura para distinguir homófonos es la conservación de formas antiguas o la reposición de letras, que retrotraen las palabras a su etimología. Por ejemplo, en el castellano del siglo XIII la palabra latina *hominem* había evolucionado a *omne* y a *ome*, prueba de que la aspiración original había desaparecido; cuando se comenzó a revisar la escritura del español, se reintrodujo la ache (el grupo /br/ es resultado de la evolución, en que no intervino el sistema de escritura) y dio *hombre*. Muchas palabras del español con ache inicial corresponden al interés por conservar el recuerdo de su etimología, lo que añade al sistema un carácter simbólico; en este caso, su prosapia latina, que no hay que soslayar, puesto que se trata de un hecho cultural.

El caso de la <v> (llamada be chica, be corta o ube) es de la misma clase. Lo mismo podemos decir del grupo consonántico <ps> en *psicología*, *psíquico*, *pseudo-*, etc. El caso de la <x> de *México* es igual; en él se percibe con mayor claridad el valor simbólico de una letra cuyo fonema correspondiente desapareció desde el siglo XVI, pero que a la ideología indigenista mexicana le recuerda el sonido de su origen náhuatl. La <ô> de *côte*, *dôme* 'catedral', *hôtel*, así como las demás vocales con acento circunflejo del sistema francés transmiten también información etimológica (en muchas de ellas, la existencia de una /s/ tras la /o/ en su voz latina originaria).

Los pocos ejemplos ofrecidos demuestran que nuestro sistema de escritura, así como los de muchas otras lenguas, no solamente están formados por sus reglas constitutivas, sino que se rigen por reglas reguladoras, determinadas por la economía del sistema (información y redundancia), para las cuales la representación morfológica, los diacríticos con que se distinguen homógrafos y la conservación de formas etimológicas son útiles instrumentos de precisión de la lengua escrita. Las reglas reguladoras, que dan su complexión al sistema, son objeto de la *ortografía*.

La *ortografía* de una lengua está formada por las reglas de complexión del sistema de escritura que determinan su *correcta escritura*. La ortografía es una construcción de la cultura de una lengua.

2. EL EFECTO DEL SISTEMA DE ESCRITURA SOBRE LA DETERMINACIÓN DE LA UNIDAD PALABRA

Varios estudiosos contemporáneos de la historia de las lenguas europeas (algunos de los cuales se citan en la bibliografía de este capítulo) han mostrado cómo la escritura se interpretaba de manera diferente en los albores de nuestras lenguas modernas, ya fuera que se tratara de las lenguas romances o de las lenguas germánicas o célticas. Para los seres humanos que tenían una lengua romance por materna, su lectura de los textos que habían heredado, en latín, se interpretaba en sus diferentes variedades dialectales (castellano, aragonés, catalán, gallego-portugués, occitano, florentino, etc.) como si se tratara de la misma lengua latina, sin preguntarse, en consecuencia, cuál era la relación entre letra y fonema (estrictamente hablando, entre letra y pronunciación, pues el fonema es una unidad lingüística apenas concebida hace poco menos de 100 años, que los hablantes no suelen poder reconocer).

Con toda seguridad sus pronunciaciones del común patrimonio verbal latino habían variado mucho durante la Edad Media, pero no se daban cuenta de ello. Los textos de que disponían —pocos, debido a la barbarie de los años posteriores a la desaparición del Imperio romano y a la agresión cristiana a la cultura pagana— se leían como si se tratara más que de escritura, de partituras, como se hace en música (en el *calpulli* azteca se hacía más o menos lo mismo con los códices jeroglíficos). En cambio, los habitantes de regiones colonizadas por Roma en el norte de Europa y en las islas británicas, cuyas lenguas maternas eran muy diferentes de las romances y, en consecuencia, no podían considerar que hubiera alguna relación entre los textos religiosos latinos y sus propias lenguas, se preguntaban con mayor perentoriedad cómo habían de leerse esos textos, con lo que la relación entre escritura y pronunciación era un problema que había que resolver.

2.1. La descodificación en fonemas y sílabas

Así que en Inglaterra, en Irlanda y en el Imperio de Carlomagno —de lengua germánica— la escritura tenía que dar lugar a una verdadera *descodificación* sonora de los textos latinos, que se hacía en silencio, pero articulando la pronunciación de sílabas y fonemas, previa a la lectura, que se solía hacer en voz alta (y es lo que haría hoy en día un niño que aprende a leer ante un texto relativamente difícil, o lo que hacemos nosotros en algunos casos, sobre todo en textos de lenguas extranjeras).

El proceso de la lectura constaba de cinco momentos: la *lectio*, en que se descifraba el texto identificando sus elementos (la *discretio* o segmentación); la *pronuntiatio*, en que se leía en voz alta; la *emendatio*, en que se corregían los errores de copia y lectura; la *enarratio*, en que se comentaban las características verbales del texto, y finalmente la *explanatio* o interpretación de su contenido. En ese momento inicial de la *discretio* era cuando se manifestaba el sistema de escritura latino como un sistema fonográfico; no es de extrañar, en consecuencia que, al comenzar a escribirse las lenguas modernas de Europa, el sistema latino se continuara y se adaptara a las características de esas lenguas (como hoy sucede con los sistemas de escritura de muchas lenguas amerindias, para los cuales la tradición escrituraria está definida por el sistema del español).

2.2. El reconocimiento de la unidad palabra

Sobre la base de los estudios de Malcolm B. Parkes, a quien hemos venido siguiendo en el párrafo anterior, se puede sostener que las gramáticas latinas heredadas en la Alta Edad Media, basadas en la teoría de las partes de la oración, es decir, ya con una concepción "lingüística" de las lenguas (aunque la lingüística, como ciencia, se inventó en el siglo XIX), suponían un reconocimiento previo de las unidades *palabra* y una percepción muy clara de la cohesión morfemática entre sus morfemas constitutivos, tanta como para poder presentar y analizar paradigmas cerrados del latín, como las declinaciones y las conjugaciones; así como para mostrar la concordancia entre sujeto y predicado, entre sustantivos y adjetivos, o las relaciones temporales entre los verbos de oraciones compuestas. La gramática latina era un instrumento muy valioso de ayuda en el momento del desciframiento, de la *discretio*.

Ese instrumento era tanto más valioso en la medida en que muchos textos estaban escritos sin espacios intermedios entre palabras; es decir, la *scriptio continua* o "escritura continua", que se encontraba en buena parte de los textos latinos heredados, imponía la necesidad de contar con instrumentos de segmentación del texto, que ayudaran a su desciframiento.

Tomemos un ejemplo para ilustrar las dificultades del desciframiento, que ofrece el mismo Parkes. Aclaremos que en la época romana no había letras minúsculas, que fueron un invento de los últimos años del Imperio. El escriba altomedieval de una copia de la *Historia de Roma* de Livio, analizó la oración:

UBIUISMAGNASPARTIADREMNAUTICAMCOLLECTA
como
ubiuis magnas parti ad rem nauticam collecta

que se interpretaría aproximadamente como: "donde muchos *partos* se habían reunido con propósitos navales", porque el escriba sólo conocía la palabra *partos* (un antiguo pueblo del actual Irán) y esa segmentación era posible. La segmentación correcta era, sin embargo:

ubi uis magna sparti ad rem nauticam collecta
"donde mucho *esparto* se había juntado con propósitos navales"

Esa clase de dificultades originó la necesidad de los copistas de los textos latinos, de abandonar la escritura continua y esforzarse por *reconocer, ante todo, unidades de denominación*, tal como las hemos venido definiendo en este libro, para distinguirlas en el continuo de la oración. Los copistas cuya lengua materna era romance, comenzaron ya fuera segmentando unidades de denominación o incluso de predicación: *sermodomini*, *cumnecessesit*, *aequesemper*, *nonconsequatur*, *aliquodrebus*, *possibileest*, como *sermo domini* 'palabra del Señor', *cum necessesit* 'necesariamente', *aeque semper* 'por siempre', *non consequatur* 'no se sigue', *aliquod rebus* 'cualquier cosa', *possibile est* 'es posible'. En cambio, afirma Parkes,

cuando los escribas germánicos y celtas escribían en sus propias lenguas, la segmentación de palabras dependía del reconocimiento de un acento principal (que dirigía la *enarratio*) y de la cohesión estrecha que reconocieran entre los morfemas que las constituyeran, con lo que, por ejemplo, escribían *isaireasber* en vez de *is aire as ber*. Pero lo que resulta claro es que la combinación del análisis gramatical, que ayudaba a reconocer, sobre todo, paradigmas de morfemas de inventario cerrado, junto con el reconocimiento de unidades de denominación, habría de conducir a una determinación final de la unidad *palabra* en la escritura. La separación en palabras mediante un espacio en blanco no se generalizó sino hasta el siglo XII d.C.

2.3. La percepción de la palabra escrita

En los ejemplos que se han ido introduciendo en el capítulo 3 y en este, en que hemos partido de pequeños textos en escritura continua para poner en práctica un análisis de cohesión morfemática, en el primer caso, y de lectura, en el segundo, nos hemos podido dar cuenta de que, aun cuando se trate de nuestra lengua materna, la longitud de las expresiones dificulta mucho el proceso de descodificación (la *discretio* de la gramática medieval).

Los pocos estudios que se han hecho de psicología de la lectura —que son, realmente, de percepción de la escritura— muestran que cuando se pide a una persona leer textos en los cuales los límites de las palabras se han suprimido o alterado, necesita fijar más la vista y mover más los ojos de un extremo al otro del texto, que cuando lee palabras separadas por espacios libres. Esos movimientos y fijaciones oculares se interpretan como indicios de que la escritura continua disminuye el campo visual, dando lugar a una menor capacidad para abarcar grandes extensiones de letras unidas unas con otras. Se puede entonces concluir que la introducción del espacio libre entre palabras modificó notablemente la capacidad lectora de los individuos y dio un valor a la escritura que no tenía en la Antigüedad ni en la Alta Edad Media.

2.4. El reconocimiento de los morfemas ligados con cohesión débil

Pero además, el apoyo que ofreció la gramática medieval al reconocimiento reflexivo, "lingüístico", de los paradigmas de morfemas de inventario cerrado, débilmente ligados a morfemas léxicos o a núcleos morfemáticos, permitió, primero, aislar esos paradigmas, y después, una vez aislados, reconocer la necesidad de darles el mismo tratamiento que a las palabras, en que predomina su carácter denominativo (los tres primeros tipos de palabra explicados en el capítulo 3), con lo cual confirmaron su carácter de palabras del tipo IV (véase el capítulo 3, § 4.1). Es decir, la palabra escrita actuó como un poderoso instrumento de reflexión sobre la lengua, que vino a actuar sobre la determinación final de la palabra oral. En ese sentido es en el que cabe considerar que la unidad palabra que manejamos hoy es una construcción de la cultura; sólo que no es una construcción caprichosa, meramente convencional, ni contingente a unas pocas lenguas, sino que se asienta sobre una existencia universal de la palabra en la oralidad, aunque las características que tenga dependan de la estructura lingüística de cada una de ellas.

El proceso de determinación final de la unidad palabra, tal como la reconocemos ahora y en muchas lenguas del mundo que tienen escritura fonográfica, no fue inmediato, sino que llevó varios siglos. Aun así, se siguen produciendo casos en que los sistemas de escritura titubean. Por ejemplo, en español *dámelo* se escribe como una sola palabra, entre espacios libres, pero *me lo das* como si fueran tres. Las reglas ortográficas del español oscilan entre considerar *asimismo* como una sola palabra o *así mismo*; se escribe *porque*, cuando se usa como conjunción y *por qué* cuando introduce una pregunta; *alrededor* y *al rededor*, etc., como dos.

De todas maneras, al reconocer la importancia que ha tenido la escritura para una lengua como el español en la determinación final de su unidad *palabra*, hemos ganado, ante todo, una integración de la escritura a los temas que conviene investigar en lingüística; hemos visto cómo la historia del español y de varias otras lenguas está profundamente imbricada con la historia de su sistema de escritura y de la lectura, y hemos demostrado, finalmente, que la unidad palabra no es irremplazable a causa de una equivocación persistente durante siglos entre las comunidades lingüísticas como la nuestra, sino que lo es porque *existe*.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

sistemas de escritura ideográfico, glotográficos y fonográficos; ícono, trazo, letra, grafía; principio fonológico; reglas constitutivas del sistema; reglas de representación; reglas reguladoras del sistema; reglas de complexión, acento prosódico, acento ortográfico; ortografía; informatividad; redundancia; economía interna de un sistema, diacrítico, homófono, homógrafo.

OBRAS CONSULTADAS

ÁVILA, Raúl, "Sólo para tus oídos", en *Lengua y cultura*, México, Trillas, 1993.

BLANCHE-BENVENISTE, Claire, «The Unit in Written and Oral Language», en Clotilde Pontecorvo (ed.), *Writing Development. An Interdisciplinary View*, Amsterdam, John Benjamins, 1997, pp. 21-45.

CATACH, Nina, *L'orthographe française. Traité théorique et pratique*, Poitiers, Nathan, 1996.

FERREIRO, Emilia, «The word out of (conceptual) context», en Clotilde Pontecorvo (ed.), *Writing Development. An Interdisciplinary View*, Amsterdam, John Benjamins, 1997, pp. 47-59.

GARRONE, Marina, "Tipografía y diseño industrial. Estudio teórico e histórico para la representación tipográfica de una lengua indígena", tesis de maestría en diseño industrial, México, UNAM, 2001.

MOSTERÍN, Jesús, *Teoría de la escritura*, Barcelona, Icaria, 1993.

PARKES, Malcolm B., *Pause and Effect. An Introduction to the History of Punctuation in the West*, Berkeley, University of California Press, 1993.

——, "La Alta Edad Media", en Guglielmo Cavallo y Roger Chartier (eds.), *Historia de la lectura en el mundo occidental*, Madrid, Taurus, 2001.

RICHAUDEAU, François, *La legibilidad. Investigaciones actuales*, Madrid, Fundación G. Sánchez Ruipérez, 1984.

SAMPSON, Geoffrey, *Sistemas de escritura. Análisis lingüístico*, Barcelona, Gedisa, 1997.

ZAMUDIO, Celia, "El papel de la escritura alfabética en la construcción del dato oral", tesis, México, El Colegio de México, 2004.

6
DE LA UNIDAD PALABRA AL VOCABLO

INTRODUCCIÓN

Una vez que hemos terminado de explicar por qué el sistema de escritura sí es determinante de la palabra en lenguas como la nuestra, que tienen una tradición milenaria de textos escritos, hemos de pasar a un último aspecto de nuestra idea social y lingüística de la palabra que resulta obvio para hablantes de culturas verbales como la hispánica, e incluso para muchos lingüistas que, aunque empeñados en negar la existencia de la palabra, no dejan sin embargo de acudir a ella en su discurso, y tampoco en sus explicaciones en el ámbito de la morfología. Tal aspecto es el de la *representación de la palabra* en la gramática y en los diccionarios, que aquí explicaremos mediante el concepto de *vocablo*.

Por lo general se piensa que la palabra *vocablo* es un sinónimo de *palabra*, que alterna uno para darle un poco de variedad al discurso. Y en efecto, si se analizan los usos de ambas palabras, *vocablo* y *palabra* varían libremente. En este capítulo no trataremos de modificar esos usos, sino de fijar la palabra *vocablo* como un útil término técnico que sirve para significar la última precisión que requiere la unidad *palabra*.

1. DE REGRESO A LA UNIDAD DE CITA

En el capítulo 2, § 3.2. habíamos visto la importancia de las *unidades de cita* para reconocer la existencia de las unidades de denominación, que constituyen la causa central de la existencia de la palabra. Explicamos cómo es una experiencia común del proceso de trabajo de un lingüista descriptivo encontrar, en las comunidades lingüísticas que estudia, que sus hablantes disponen de ciertas formas del contenido de la palabra para poderla citar, en casos en que hacen preguntas sobre su significado, responden preguntas que las involucran o introducen observaciones acerca de ellas.

Dimos algunos ejemplos del náhuatl de Ocotal Texizapan, Veracruz, en que se comprueba que los verbos se citan en tercera persona del singular del presente: [motalowa] 'corre', [nehnemi] 'camina', [pata:ni] 'vuela', o que los sustantivos se citan, bien sin morfemas de género y número [ikpál] 'silla', [ilwika'] 'cielo', o bien con un morfema "posesivo" [noma] 'mi mano', [nohyish] 'mi ojo', etc.[1] En español, estamos acostumbrados a citar los verbos en su forma infinitiva y los sustantivos en su forma masculina y singular. En este capítulo buscaremos explicar a qué se debe este fenómeno y cuál es su importancia para completar nuestra comprensión de la unidad palabra.

1.1. El papel de la reflexión en la determinación de la palabra

Es fácil colegir que una persona que es capaz de encontrar una forma de contenido que represente a una palabra de su vocabulario cada vez que la quiere mencionar o citar fuera de contexto, lo puede hacer porque en su comunidad lingüística se ha producido una capacidad reflexiva que permite a sus miembros observar su propia lengua y llevar a cabo un análisis de sus signos semejante al que hemos venido siguiendo en este libro, aunque de carácter distinto, pues no llega a convertirse en una lingüística. De ese análisis tradicional y propio de cada comunidad lingüística resultan varios conocimientos compartidos por la comunidad acerca de su propia lengua y transmitidos mediante su educación. Que existe un análisis de la propia lengua en todas las comunidades lingüísticas, lo demuestra, por ejemplo, el fenómeno de la rima, que es tan antiguo como la memoria humana y forma parte central del canto y de la poesía.

Poder rimar dos palabras implica darse cuenta de la semejanza sonora que hay entre ellas; si pensamos en los metros de la poesía, ya sea los basados en la cantidad silábica o en el número de sílabas y los juegos acentuales, vemos que desde la más remota antigüedad ha habido reflexión acerca de la forma de la expresión de los signos de las lenguas.

[1] En este capítulo los ejemplos del náhuatl proceden de dos fuentes: de encuestas realizadas por estudiantes en diferentes localidades de lengua náhuatl, pero diferentes dialectos, y del náhuatl escrito que se encuentra en los diccionarios. Los ejemplos de la primera clase se citan en transcripción fonológica entre corchetes; los de la segunda, con el sistema de escritura tradicional de los estudiosos del náhuatl, en cursivas.

Quien tiene contacto con niños de tres o cuatro años de edad en adelante, puede comprobar que la reflexión sobre su propia lengua es un ejercicio constante, característico de los procesos de disociación de objetos en las acciones verbales y de objetivación de los signos que los significan.

No es por eso extraño que todas las comunidades lingüísticas reconozcan las características de sus palabras —sobre todo de las que corresponden claramente a unidades de denominación, como las que hemos clasificado como de tipos I, II y III en el tercer capítulo de este libro— y lleguen a elegir, de todas las formas que ofrece un determinado paradigma flexional, derivacional o de declinación, una forma en particular que se convierta en la unidad de cita de todas las palabras que conforman el paradigma.

La capacidad de reflexionar sobre la propia lengua se ve potenciada por la escritura, pues ésta da permanencia a lo que habría pasado fugazmente en el habla y permite considerarlo con lentitud y someterlo a observación y análisis. No de otra manera se habría producido nuestra tradición gramatical greco-latina, ni habrían aparecido obras como los diccionarios, dedicadas al vocabulario de cada lengua. Como vimos en el capítulo anterior, cuando la escritura ayuda a segmentar morfemas constitutivos de la palabra, a reconocer su pertenencia a determinados paradigmas o a separar otros cuya cohesión con la palabra o no existe o es muy débil, la reflexión sobre la lengua misma aumenta y las características de la unidad palabra se vuelven más precisas.

Así que la existencia de unidades de cita en cualquier lengua es una muestra de la reflexión que hace la comunidad lingüística acerca de su propia lengua, transmitida tradicionalmente de generación en generación, y una muestra de la conveniencia de distinguir entre *palabra* y *vocablo*, atribuyéndole a éste un carácter representativo, que explicaremos más adelante.

1.2. Características de las unidades de cita

Veamos ahora algunas de las características que tienen esas unidades de cita, basándonos en los pocos documentos que hemos encontrado acerca de ellas, pues desgraciadamente no han sido objeto de suficiente atención por parte de la lingüística descriptiva.

a] Puesto que la función de la unidad de cita es representar un signo cuando se le aísla en el discurso y se le menciona, toda unidad de cita está constituida por un número entero de sílabas y se ajusta a la función demarcativa de los fonemas, cuando es el caso de que haya tales funciones en cada lengua determinada.

b] La unidad de cita consiste, desde el punto de vista morfológico, de la menor cantidad de morfemas flexionales o derivativos que sea necesaria para que la palabra tenga significado, pero que a la vez no añadan significado específico a ella. Por ejemplo, si tenemos en náhuatl [kalli] 'casa', [nokal] 'mi casa' y [nokalwan] 'mis casas', puesto que los morfemas posesivos especifican de qué 'casa' se trata, se prefiere la forma que no los tiene, [kalli], como unidad de cita; en los tres casos un nahuatlato afirmará que se trata de palabras, debido a la cohesión estrecha que se produce entre la raíz *-kal-* y los morfemas que se agrupan con ella. Como consecuencia gramatical de ello, por la reflexión que hace el hablante se puede segmentar un morfema *-li* en oposición con *-wan* o su desaparición en [nokal] (aunque con el morfema posesivo prefijado); la gramática náhuatl ha considerado ese morfema *-li* como "absolutivo", es decir, que se utiliza para citar la palabra sin declinación. En *coatl* 'serpiente' encontramos el morfema *-tl*, también absolutivo, por lo que, al compararlo otros morfemas con la misma función, como *-tli* (*tochtli* 'conejo') e *-in* (*michin* 'gato'), reconocemos un paradigma de morfemas absolutivos nahuas, que sirven para mencionar sus unidades de cita de los sustantivos. A la vez, de las tres formas [kalli, nokal y nokalwan], se elige [kalli] como representante de la palabra en la mención, que descontextualiza el signo y lo aísla.

Si pensamos en el español sucede algo semejante: la unidad de cita de un sustantivo es su forma con morfemas masculino y singular, si se trata de palabras como *hermano*, *niño*, *perro*, *gato*, etc., pues tanto los morfemas de plural como los de género femenino especifican el significado. Por eso en español, el género masculino se considera neutral. Tomemos el siguiente ejemplo: si se dice "en este salón hay cuatro alumnas", se entenderá que hay cuatro mujeres; si se dice que "hay cuatro alumnos", no se especifica si se trata sólo de hombres o de hombres y mujeres. A esta propiedad del género masculino en español (¡ojo: no confundir con el *sexo* masculino!) se llama "masculino genérico". Es verdad que a veces el masculino genérico del español lleva a soslayar la

participación o incluso los derechos de las mujeres, como justamente señalan los defensores de los derechos humanos, pero en esos casos se puede especificar (sin caer en el ridículo en que han caído varios políticos). Cuando un sustantivo es de género femenino, naturalmente la unidad de cita es la femenina singular, como *foca* (por eso hay que decir "foca macho") *tela*, *cabeza*, etc. Según la clase de sustantivos de que se trate, se determina la unidad de cita correspondiente.

En cuanto a los verbos sucede lo mismo. Ya vimos en el segundo capítulo que la unidad de cita de los verbos nahuas consiste en su forma de tercera persona del singular del presente, pues es la menos específica, tanto morfológicamente como en cuanto a su significado. Algo semejante ocurre con muchas lenguas amerindias, como el tzeltal. En español, ese papel lo tiene la forma infinitiva, que está desprovista de morfemas de persona, de número, de tiempo y de modo.

c] En una lengua hay varias unidades de cita, según los tipos de palabra que haya en ella; cuando se trata de palabras de tipo IV —es decir, de conjunciones, preposiciones, artículos, pronombres, etc.—, cuya identificación es resultado de la segmentación morfológica y la medida de la cohesión que tengan con sus correspondientes núcleos morfemáticos, las comunidades lingüísticas sin tradición de escritura titubean y les llega a costar trabajo tanto citarlas como segmentarlas, y así lo comprueban todos los lingüistas que, por ese motivo, ponen en duda la existencia de la unidad palabra en todas las lenguas.

1.3. El papel de la unidad de cita

La unidad de cita es una forma representante de la palabra, en la que se abstrae la variedad morfológica con que aparece usada en el discurso, para poderla mencionar con facilidad y para poderla objetivar como tema de reflexión. En cuanto tal, estrictamente hablando, es la representante del paradigma que forman la raíz o el núcleo morfemático de la palabra y los morfemas ligados en cohesión estrecha con ellos, sobre la base de la unidad de denominación.

Por ejemplo, si ya hemos dicho que la forma de cita de los verbos en español es su infinitivo, podemos darnos cuenta de que éste representa a todas las flexiones de modo, tiempo, número y persona de cada ver-

bo. Así, *cantar* es la unidad de cita de *canto*, *cantas*, *canta*, *cantamos*, *...canté*, *cantaste*, *cantaron*, *...cante*, *cantes*, *cantemos*, ... etc. En cuanto a los sustantivos, *gato* es la unidad de cita representante de *gatos*, *gata*, *gatas*, *gatitos*, etcétera.

Si cada uno de los signos que forman los paradigmas de conjugación de un verbo, o las flexiones y derivaciones de un sustantivo *son palabras*, identificadas mediante los criterios y procedimientos explicados en este libro, llamaremos a sus correspondientes unidades de cita, que también son palabras, *vocablos*.

Un vocablo es, entonces, una unidad de cita que representa todo el paradigma de flexiones, derivaciones o conjugaciones que se forman a partir de una raíz o de un núcleo morfemático y que se reconocen como palabras.

2. PALABRA Y VOCABLO

Se puede ahora establecer la convención terminológica de llamar *palabra*, primero, a toda unidad de denominación cuya forma de la expresión consta de un número entero de sílabas y cuyos límites pueden estar definidos por la función demarcativa de algunos de sus fonemas, y cuya forma del contenido consta de un morfema léxico, ligado en cohesión estrecha con otro(s) morfema(s) léxicos para constituir núcleos morfemáticos, y con morfemas pertenecientes a paradigmas cerrados de carácter flexional, conjugacional o derivativo; segundo, a todo signo deslindado posteriormente entre las unidades de denominación, mediante una cohesión débil o inexistente, perteneciente a paradigmas cerrados de carácter sintáctico.

Se puede, en cambio, llamar *vocablo* a toda forma del contenido que se adopte como representante de un paradigma completo de flexión, de conjugación o de derivación correspondiente a una unidad de denominación.

Al hacerlo así, hemos introducido una convención terminológica que redefine las palabras *palabra* y *vocablo* como términos técnicos de la lexicología; *palabra* será ahora un término de observación y descripción,

correspondiente a un fenómeno real de las lenguas; *vocablo* será un término de descripción y primera teorización, que sirve como instrumento abstracto de la investigación lingüística y del método lexicográfico.

2. LA CONSTITUCIÓN HISTÓRICA DEL VOCABLO

Es evidente que el *vocablo* es una unidad construida por la reflexión sobre la lengua y, por lo tanto, de carácter secundario en comparación con la *palabra*. Varias investigaciones de la historia de los diccionarios en Occidente revelan que el proceso de abstracción que llevó a fijar vocablos como representantes de palabras duró muchos años, y se vio impulsado por el pensamiento gramatical greco-latino y por la necesidad de la lexicografía de utilizar alguna forma representante de la palabra en la entrada de sus artículos lexicográficos.

Esa forma no sólo debía corresponder a la unidad de cita que, como hemos visto, es producto de una reflexión intrínseca a las culturas verbales de las lenguas del mundo, independientemente de si tienen o no sistema de escritura, sino que debía facilitar el reconocimiento de las demás formas flexionadas, conjugadas o derivadas que se pueden producir mediante la unión de una raíz y diferentes morfemas en una unidad de denominación.

Tomemos por ejemplo el caso del latín: de la declinación de *rosa*, *rosæ*, *rosam*, *rosarum* (citamos sólo las formas diferentes) se seleccionó como unidad de cita la forma nominativa, que no tiene morfemas especificadores. Se impuso como *convención lexicográfica* considerar la palabra *rosa* como vocablo y además como representante del paradigma de la primera declinación (seguida de la forma del genitivo, *rosæ*, para aclarar la declinación). Del mismo modo, se seleccionaron palabras correspondientes a cada una de los otros cuatro paradigmas de declinación latinos como vocablos, representantes tanto de la palabra como de la declinación. En cuanto al verbo, la lexicografía del latín —que es una invención posterior a la desaparición del Imperio romano— no pudo ofrecer una solo vocablo como representante de su conjugación, porque no permitía recuperar inequívocamente de él todas las formas conjugadas, sino que se vio en la necesidad de citar los verbos en los diccionarios mediante los *vocablos* representantes de la primera persona del singular

del presente (*amo*), de la segunda persona (*amas*), el infinitivo (*amare*), la primera persona del singular del pretérito perfecto (*amavi*) y la forma del supino (*amatum*). Un verbo tan irregular como *fero, fer, ferre, tuli, latum* demuestra la imposibilidad de reducir esos cinco vocablos a uno solo. Lo mismo sucede, en realidad, con el verbo *ir* en español, en que su infinitivo no basta para representar los vocablos *voy, iba, fui* e *iré*.

El vocablo, en consecuencia, es un producto de la cultura de la lengua, construido mediante el análisis gramatical y las necesidades de los métodos de elaboración de diccionarios.

Una vez que podemos abstraer las palabras como vocablos, hemos ganado una forma sencilla de referirnos a ellas como miembros de paradigmas de flexión, de conjugación y de derivación, que facilita mucho el discurso lingüístico.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

vocablo; reflexión sobre la lengua; convención terminológica; lenguajes de observación; de descripción y de teorización.

OBRAS CONSULTADAS

HEGER, Klaus, *Monem, Wort, Satz und Text*, Tubinga, Max Niemeyer Verlag, 1971.
LYONS, John, *Introducción en la lingüística teórica*, Barcelona, Teide, 1977.

SEGUNDA PARTE
EL LÉXICO Y SU ESTUDIO

INTRODUCCIÓN:
EL LÉXICO Y SUS ABORDAJES

La existencia real de la unidad *palabra* en una lengua es lo que posibilita y da lugar a la consideración de su conjunto, que llamamos *léxico*. Si la *palabra* se nos hace presente en el discurso y en el texto, y como hablantes de una lengua somos capaces de convertirla en objeto de reflexión, el *léxico* sólo se nos manifiesta cuando pensamos en la clase de unidades verbales que llamamos *palabras* y nuestra memoria nos permite recordar muchas de ellas. Cuando, por ejemplo, pensamos en el vocablo *gato* y vienen a nuestra mente otros con los que hablamos de los animales domésticos, como *perro*, *caballo*, *burro*, *gallina*, *cabra*, etc., ese conjunto forma un *léxico*. Si seguimos haciendo el mismo ejercicio, dirigiendo nuestra atención a cuanto ámbito de nuestra experiencia se nos ocurra, iremos sumando vocablo tras vocablo a ese léxico que descubrimos existente en nuestra memoria. El léxico se nos manifiesta en primera instancia como un fenómeno de la memoria de cada individuo. Pero en cuanto se va alojando en ella a lo largo de la vida, de manera ilimitada, como parte de la lengua que cada quien recibe de su comunidad lingüística, no es un léxico privado, sino sólo aquella parte del gran acervo de la lengua histórica que se recibe durante el aprendizaje de la lengua y su consecuente educación. Sólo a partir de ese hecho es como podemos, después, poner nuestra atención en ese conjunto de vocablos, en esa clase de signos lingüísticos que son los vocablos, que descubrimos en la lengua compartida por todos los miembros de una comunidad lingüística y que llamamos "léxico de una lengua".

La escritura, que es un instrumento no sólo de la comunicación entre individuos que no están presentes, sino también de conservación de discursos escuchados o directamente escritos como textos, vuelve a manifestarse, al igual que en el reconocimiento de la palabra, como una gran ayuda para trascender la memoria individual y construir una memoria colectiva del léxico. La aparición de los diccionarios, que son un

efecto de la escritura, ha ampliado de una manera vertiginosa la memoria colectiva del léxico, hasta el punto de superar las memorias individuales y ofrecernos grandes acervos léxicos, que sin lugar a dudas son el elemento central de la riqueza de las lenguas y de la posibilidad humana de conservar la memoria de sus experiencias compartidas.

Como fenómeno individual y social, el léxico tiene tres dimensiones: una lingüística, una cognoscitiva (psicológica) y una neurológica. La dimensión lingüística es necesariamente anterior a las otras dos, pues tanto la psicología como la neurología dependen de la manifestación del fenómeno lingüístico para poder comenzar a estudiarlo. Dicho de otra manera, son la palabra y el léxico los *observables iniciales* de esas dos ciencias; son los fenómenos materiales concretos que inducen a su exploración en el ámbito de la cognición y del funcionamiento de la memoria en el cerebro. De no ser así, el estudio de la cognición se extraviaría en una multitud de respuestas conductuales del ser humano, a las que no se les podría encontrar alguna clase de organización. Lo mismo sucedería en la investigación neurológica: entre los millones de respuestas de la corteza cerebral a los estímulos sensoriales que se producen a cada instante de la vida, no se podría saber nunca cuáles de ellos corresponden a las palabras y a la memoria de su conjunto. Toca, por lo tanto, a la lexicología definir las características y las propiedades del léxico, en tanto que corresponde a la psicología y a la neurología contrastar los conocimientos que les ofrece la lexicología con los datos que obtienen con sus propias hipótesis y sus métodos, para así llegar a conocer la existencia del léxico en la memoria y en el cerebro.

Reconocer estas tres dimensiones del léxico es importante por dos motivos: el primero, porque abren a la investigación lingüística un conjunto de fenómenos de memoria que, hasta la fecha, sólo se han considerado en su aspecto utilitario, para la elaboración de métodos de enseñanza de lenguas y para la preparación de pruebas de diagnóstico de diferentes problemas neurológicos y audiológicos, en vez de darles su lugar en la comprensión integral del fenómeno del léxico; el segundo, porque consecuentemente con lo anterior, permite integrar la existencia del léxico al conjunto de cuestiones que plantea la búsqueda de las características sistemáticas de las lenguas, sacando a la lingüística de los muy estrechos estudios formalistas contemporáneos, para los que el léxico sólo tiene valor de *lista de vocablos*.

En esta segunda parte del *Curso* nos ocuparemos, por ello, de la dimensión lingüística del léxico, pero a la vez crearemos los lazos mínimos necesarios entre la lexicología y los intereses de la psicología y la neurología. Procederemos exponiendo, primero, algunos de los métodos de recolección de datos léxicos de que se puede disponer hoy en día; después, métodos cuantitativos de estudio del léxico; continuaremos con un capítulo dedicado al estudio cualitativo del significado de los vocablos en un léxico (no tocaremos los estudios de la forma del contenido del vocablo, que corresponden a una morfología, con sus apartados de morfo-fonología y formación de palabras); en otro más trataremos la manera en que el léxico manifiesta diferentes fenómenos normativos, y terminaremos dedicando nuestra atención a los estudios etimológicos del vocabulario, tan importantes para la cultura de la lengua

7
LA RECOLECCIÓN DE DATOS LÉXICOS

INTRODUCCIÓN

A diferencia de los demás elementos que conforman una lengua, como los fonemas, los morfemas y los patrones oracionales, que puede uno encontrar en una relativamente pequeña colección de grabaciones, de textos o de respuestas a un cuestionario, el léxico de una persona consta de más de mil vocablos, y el léxico de una lengua histórica de una cantidad indeterminada e indeterminable de vocablos. Tal indeterminación se debe al hecho de que, de todos los elementos de una lengua, son las palabras las que se relacionan directamente con la experiencia del mundo y de la vida que, por naturaleza, varían en el tiempo y en el espacio y dependen de la novedad de cada experiencia humana socialmente compartida.

El léxico de una lengua es indeterminable por ser ilimitado, es decir, aumenta, varía o desaparece de manera ilimitada. Podemos saber con cierta aproximación qué cantidad de vocablos hacen falta para hablar una lengua (el llamado *vocabulario fundamental*), pero no de cuántos vocablos consta una lengua histórica, por más que el público crea que la cantidad de vocablos reunidos en un diccionario corresponde al tamaño del léxico de la lengua, o crea las falsedades que acostumbran aparecer de vez en cuando en la prensa, según las cuales hay lenguas "más ricas" que otras.

Podemos afirmar que todas las lenguas del mundo disponen de un léxico ilimitado; a diferencia de sus diccionarios, cuya limitación (en "pobreza" o en "riqueza") proviene de sus fuentes, de la cantidad de trabajo que dedican sus autores a encontrarlos y registrarlos, y de la calidad del acervo que han podido ir acumulando.

Dado el carácter ilimitado del léxico, cada vez que nos planteemos estudiarlo tendremos que planear previamente y con cuidado cómo recolec-

En el caso de las lenguas amerindias, cuyo desarrollo cultural se vio interrumpido por la conquista, la colonización y la imposición de una sola lengua sobre ellas, su léxico no parece ser bastante para hablar del mundo contemporáneo, lo que requeriría procesos de enriquecimiento léxico a base de una neología cuidadosamente construida, como se hizo, por ejemplo, con el hebreo moderno.

tarlo. Su recolección dependerá del objetivo que persigamos: si queremos estudiar el léxico de una lengua histórica como tal, tendremos que construir los instrumentos necesarios para encontrarlo, registrarlo, documentarlo y manejarlo; si queremos estudiar el léxico de cierto grupo social, tendremos que definir las características de ese grupo para diseñar la manera de obtenerlo; si queremos preparar un léxico para la enseñanza de una lengua o para un cuestionario audiológico o neurológico, tendremos que saber previamente qué características del léxico son pertinentes para la enseñanza o para ese cuestionario, ya en cuanto a su forma del contenido, ya en cuanto a su sustancia del contenido.

1. CONCEPTOS BÁSICOS

Conviene distinguir, por lo tanto, cuatro conjuntos básicos del léxico: en primer lugar, como ya se adelantó, aquel léxico mínimo que permite hablar una lengua en situaciones de inmediata necesidad comunicativa, sin pretender referirse a temas específicos, constituye el *vocabulario fundamental*.

1.1. El vocabulario fundamental

Los primeros estudios de vocabulario fundamental no fueron obra de lingüistas, sino de los inventores de sistemas taquigráficos, qué necesitaban saber qué vocablos eran los imprescindibles en cada lengua, para asignarles a cada uno de ellos un trazo que los representara —simplificando, en consecuencia, la complejidad formal de la palabra a un solo trazo corto y fácilmente perceptible— y redujera el trabajo taquigráfico, a la vez que aumentara la velocidad de la codificación. Otra clase de estudio del vocabulario fundamental se obtiene indirectamente de las listas de ideogramas, que las tradiciones pedagógicas china y japonesa llegaron a determinar como el mínimo necesario para que un alumno de escuela elemental fuera

capaz de aprender a leer y escribir el chino o los *kanji* (ideogramas de la escritura china, adoptados por la escritura japonesa). En una y otra clase de estudios, el vocabulario fundamental resulta estar compuesto por un poco más de 1 000 vocablos, hasta cerca de 2 000, según la lengua de que se trate, pero nunca es mayor. Es decir, podemos suponer que hay una constante mínima del léxico con que se puede hablar una lengua, y que esa constante es de aproximadamente 2 000 vocablos. En el capítulo siguiente veremos cómo se reúne y se calcula el tamaño del vocabulario fundamental.

En la investigación que hicimos para construir el *Diccionario del español de México*, determinamos como vocabulario fundamental del español mexicano de tradición culta un conjunto de 1 451 vocablos; el vocabulario fundamental común a esta tradición y a la tradición popular resultó de 1 131 vocablos. (Acerca de los conceptos de "tradición culta" y "tradición popular", véase el capítulo 9.)

1.2. El vocabulario activo

El vocabulario fundamental es el que nos permite hablar una lengua en condiciones mínimas de comunicación; es decir, es un vocabulario de actuación primaria. Pero el hablante de una lengua ya ha aprendido en sus primeros años de vida y antes de entrar a la escuela muchos más vocablos que los correspondientes al vocabulario fundamental; aprendió a hablar, quizá, de alimentos, de animales, de juguetes, de parientes, de juegos, de emociones, de religión, etc. Con ese vocabulario, que supera la cantidad mínima necesaria y llega a constar de algunos miles de vocablos, el individuo puede conversar, dar expresión a sus más particulares experiencias, dar instrucciones, leer libros y ponerse de acuerdo con los demás para llevar a cabo muchas acciones. A ese vocabulario lo llamaremos *activo*. Es un vocabulario que domina el hablante, que maneja espontáneamente y con el que se desempeña en su vida cotidiana.

No es posible llegar a saber cuál es el vocabulario activo de cada persona, pues es imposible diseñar un instrumento de investigación que nos permita obtener de ella todo el léxico que maneja con facilidad; pero sí es posible suponer, gracias a la experiencia de algunos lexicógrafos, que contrastan diariamente sus conocimientos del vocabulario con el léxico documentado de la lengua en que trabajan, que el vocabulario activo de una persona va de los 2 000 a los 8 o 9 000 vocablos de la lengua histórica.

1.3. El vocabulario pasivo

En cambio, toda persona, durante su educación y durante su vida, va escuchando y leyendo muchos más vocablos de los que llega a utilizar activamente, que quedan grabados en su memoria y le permiten comprender nuevas conversaciones, nuevos textos, con poco esfuerzo. Ese vocabulario aprendido, que no se utiliza de manera activa, es el *vocabulario pasivo*.

Si es imposible explorar hasta el límite un vocabulario activo, es igualmente imposible explorar el vocabulario pasivo de una persona. Éste puede llegar a constar de muchos miles de vocablos, según los intereses de cada quien, según sus estudios o según su ocupación. Así por ejemplo, una costurera conoce muchos vocablos relacionados con las telas, los hilos, las puntadas, los cortes, las proporciones de la ropa, etc., que utiliza activamente, y a la vez puede comprender varios vocablos relacionados con medios de transporte o con la economía, como *teleférico*, *autovía*, *ferry*, *insumo*, *déficit*, etc. que, sin embargo, no maneja; en cambio, un piloto de avión tendrá un vocabulario activo en el que el léxico de la aviación sea muy grande, y uno pasivo de ingeniería, aerodinámica, resistencia de materiales, comunicaciones, etc. que le permitan comprender muchas cosas, aunque no lo utilice; su conocimiento del vocabulario de la costura será mínimo. Un filólogo o un lingüista, cuya profesión toma como núcleo las lenguas, tendrá un vocabulario pasivo mucho más extenso que los dos anteriores, aunque su vocabulario activo, quizá grande, sea mucho menor que el pasivo.

De este hecho se puede sacar una recomendación: nadie debiera poder afirmar, en relación con una palabra que oye por primera vez, que "no existe" o que "no se usa"; lo único que se puede afirmar es que uno no la conoce o no la usa. Esta recomendación es especialmente importante para el comportamiento ético de un lingüista.

De todas maneras, el conocimiento activo y pasivo del vocabulario en cada persona es diferente y nunca es equivalente al léxico de su propia lengua. En relación con el léxico de la lengua histórica, todo conocimiento individual es parcial y fragmentario.

1.4. El vocabulario disponible

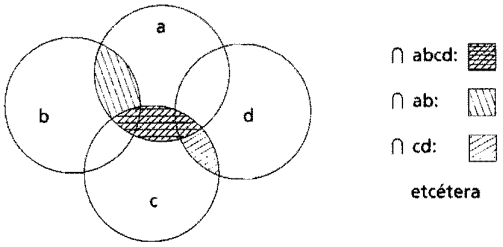
También es imposible trazar una frontera entre el vocabulario activo y el pasivo de una persona. Primero, porque durante toda la vida aprende uno nuevas palabras y, en consecuencia, uno y otro conjunto se modifican día a día; pero también porque la memoria humana es capaz de hacer pasar vocablos de un conjunto al otro. Imaginemos el siguiente ejemplo: rara vez hablamos de terremotos, salvo cuando se produce uno. Para hablar de ellos, disponemos espontáneamente al menos de dos vocablos: *temblor* y *terremoto*. Suponemos que ambos forman parte de nuestro vocabulario activo. Pero basta con leer la prensa o escuchar una noticia por radio, para que se actualicen en nuestra memoria vocablos que quizá formaban parte de nuestro conocimiento pasivo: *sismo*, *escala de Mercalli*, *escala de Richter*, *onda*, *réplica*, *volcán*, *flujo piroclástico*, *erupción*, *lava*, etc. y, durante nuestra conversación o en un texto que estemos escribiendo, comenzaremos a manejarlos con facilidad. En ese momento, pasan a formar parte de nuestro vocabulario activo.

Este es un fenómeno de la memoria, que puede estar influido por muchos factores, como la edad, el sexo, la educación, el conocimiento recibido de la tradición verbal, la especialidad de cada quien y hasta, quizá, por flujos represivos del inconsciente, pero manifiesta la disponibilidad de vocablos en la memoria individual. Cuando se investiga este fenómeno en cada persona, se puede determinar un *vocabulario disponible*. En el capítulo siguiente veremos cómo se hace esta clase de investigación.

2. DEL LÉXICO INDIVIDUAL AL LÉXICO COMPARTIDO POR UN GRUPO Y AL LÉXICO DE LA LENGUA HISTÓRICA

Hechas las distinciones anteriores, que toman como necesario punto de partida el conocimiento individual de las personas, ahora podemos preguntarnos cómo llegar a conocer el léxico de una lengua histórica. Lo primero que podemos reconocer es que toda conversación entre personas que hayamos podido grabar como dato de nuestra investigación, así como todo texto, ya sea de autor conocido o anónimo que hayamos recogido con esa finalidad, revelan un vocabulario activo; sólo en relación con otras personas, oyentes o lectoras, parte del vocabulario de esas fuentes se puede convertir en pasivo, si no hacen uso de él. Cuando estudiamos grabaciones y textos para llegar al léxico de la lengua histórica, todo el vocabulario que encontremos será, en principio, pasivo en su conjunto, aunque muchos subconjuntos suyos sean activos para individuos determinados.

Cada conjunto de vocablos obtenidos de fuentes individuales se intersecta o traslapa en diferentes medidas con los otros conjuntos. Si representamos cada fuente con una letra del alfabeto, podemos esquematizar esta idea de la siguiente manera:



La intersección ⋂(a,b,c,d) nos da todo el *léxico común* a esos cuatro vocabularios individuales, pero deja fuera de consideración todo el léxico individual no compartido por los otros tres conjuntos. Hay intersecciones entre el léxico de a y b, de a y c, de a y d; entre el de b y c, y b y d; entre el de c y d; etc. La suma o adyunción ⋃(a,b,c,d), en cambio, de los cuatro conjuntos, nos da un léxico total, y, como no hay léxico privado (todo aparente signo lingüístico que emita una persona no es signo mientras no sea comprendido por los demás), podemos considerar que obtenemos un acervo exhaustivo de ellos.

Para estudiar el vocabulario de un grupo social, por ejemplo, de los pescadores de río en Tlacotalpan, de los artesanos del barro en el valle de Oaxaca, de las educadoras en un jardín de niños, de los jóvenes de un barrio, etc., nos sirve la intersección de sus vocabularios individuales, es decir, su *léxico común*; en cambio, para estudiar el léxico de la lengua histórica, ya sea a lo largo de la historia o en diferentes estadios de su existencia, nos sirve la suma de todos los vocabularios que hayamos podido reunir.

2.1. Los *hapax legomena*

Heredamos esta expresión griega de *hapax legomenon* (en singular), que quiere decir 'dicho una sola vez', de los estudios filológicos dedicados a la obra del historiador griego Jenofonte, en la cual se encuentran muchas expresiones que aparecen una sola vez y que no se han encontrado en el resto del acervo de textos griegos clásicos, por lo que son muy raras y crean dificultades a veces insalvables de interpretación, pues aunque puede tratarse de verdaderas palabras, también pueden ser simples errores de un copista. En un registro de palabras en textos o en conversaciones con un solo individuo es posible encontrar algunas que desconozcamos y que no podamos comprobar en otros registros; podemos hasta dudar de su carácter de signo, por lo que tendremos que considerarlas *hapax* mientras no podamos encontrar al menos una aparición suya más, en otras fuentes. De la suma de vocablos recogidos en el ejemplo anterior, tendremos que restar, consecuentemente, los *hapax*.

2.2. La formación de un corpus de datos

Cuando buscamos el léxico de una lengua histórica tenemos que proceder sumando conjuntos de vocablos de todas las fuentes posibles; en cambio, cuando queremos encontrar un léxico característico de cierto tema, cierto grupo social, cierta época, etc., tenemos que diseñar un método de recolección de datos léxicos, que nos dé suficiente seguridad de que lo que encontremos corresponda a aquello que buscamos.

Dijimos al comienzo de este capítulo que lo característico del léxico, en comparación con los demás elementos de una lengua, es la inexistencia de límites. Cuando se trata de la lengua histórica, en efecto, no podemos saber de cuántos vocablos consta a lo largo de su historia; tampoco de cuántos vocablos consta en alguna de sus variedades regionales, como el español en México, en España, en Colombia, en el occidente de México, en el español murciano o en el de la sabana de Bogotá. Lo mismo podemos decir de las variedades del tzeltal en Chiapas o del zoque en Oaxaca y Chiapas.

En relación con el léxico de una lengua histórica nos interesa cada uno de sus vocablos, pues cada uno tiene un significado particular, que lo vuelve diferente de todos los demás y de aquellos con los que tiene alguna cercanía, algún parentesco etimológico o alguna afinidad. Tenemos en ese caso un interés *cualitativo* por el léxico. En cambio, si queremos reunir el léxico de una lengua para someterlo a un estudio de sus formas de contenido (su conformación morfológica, las mane-

ras en que se relaciona con la sintaxis, las reglas con que forma nuevas palabras, etc.) o a una investigación de su uso en cierta clase de discursos, en cierta sociedad, entre los miembros de cierto grupo, no nos será necesario registrarlo de manera exhaustiva, sino sólo llegar a determinar un conjunto de vocablos que valga como *representativo* de él, como para poder sacar conclusiones generales acerca de sus características. En ese caso tenemos un interés *cuantitativo* por el léxico, además de cualitativo.

2.2.1. El problema de la representatividad

El concepto de *representatividad* de un conjunto de datos léxicos en relación con el total del vocabulario de una lengua o de una de sus variedades presenta un problema serio, que debe uno tomar en cuenta para resolverlo de la mejor manera posible. Para ilustrar este problema tomemos un ejemplo que nos sirva de contraste: durante los períodos electorales en las democracias modernas, desde el momento en que los partidos políticos seleccionan a sus candidatos a puestos de elección popular, tanto a ellos como a los ciudadanos les interesa tener una idea de las posibilidades de triunfo de cada uno. Para medir esas posibilidades, las agencias de prospección política conforman una *muestra* de electores, que constituye un porcentaje de la población total con derecho a voto, pues consultar a todos equivaldría a la votación misma.

Se busca que la muestra sea representativa del total y para ello se toma en cuenta el número total de electores —que se conoce gracias a las listas del padrón electoral—, del cual se selecciona, sobre la base de un estudio de probabilidades, un porcentaje de ellos, que esté formado por hombres y mujeres, de diferentes edades, de diferentes ocupaciones, etc. De ese modo, se trata de impedir que la muestra tenga sesgos (por ejemplo, si se consultara sólo a las personas que tienen coche). Los resultados que se obtengan del estudio de esa muestra se extrapolarán, mediante diversos métodos estadísticos, a la población total, lo que dará por resultado una idea muy precisa de las preferencias de voto de los electores, aunque siempre con cierto margen de error, que también puede calcularse.

El problema de la lingüística es que, tratándose del léxico, como no podemos determinar la "población léxica" total, pues hemos dicho que es ilimitada y por eso indeterminada, tampoco podemos seleccionar un porcentaje adecuado de vocablos que la representen.

La única manera de lograr que una muestra léxica sea relativamente representativa en relación con el vocabulario de la lengua histórica es reunir textos o transcripciones de emisiones orales, cuidando que, si se trata de textos, esté compuesta de una rica variedad, como novelas, cuentos, ensayos, cartas, obras de teatro, etc., y si se trata de discursos grabados, que sean de conversaciones espontáneas, alocuciones políticas o religiosas, actos cívicos, etc. en una cantidad que resulte *suficiente*. Pero, ¿qué es suficiente? La suficiencia de un corpus depende, primero, de que hayamos tratado de eliminar posibles sesgos en la muestra, asegurándonos de que la selección de los datos haya sido *aleatoria*; luego, de la *variedad* que le hayamos dado a nuestra recolección; después, de su *cantidad*; por último y de manera más importante, de la *riqueza léxica* que el propio corpus nos va mostrando durante su análisis; es decir, la suficiencia parece ser un *resultado* que no se puede prever, y no una medida predeterminada.

Sin embargo, la experiencia acumulada por la lingüística en el estudio de corpus de datos léxicos nos permite enfrentar este problema con buenas posibilidades de éxito, aun cuando no podamos dar unas proporciones seguras para planear un corpus.

2.2.2. Ocurrencia, tipo y vocablo

Para comprender mejor lo antes enunciado, hagamos tres convenciones terminológicas que nos resultarán muy útiles:

a] cada aparición de una palabra en un texto será una *ocurrencia*; por ejemplo, en el siguiente fragmento del romance de "El enamorado y la muerte", contamos 20 ocurrencias:

"Ábreme la puerta, blanca,/ ábreme la puerta, niña./¿La puerta cómo he de abrirte/ si la hora no es convenida?"

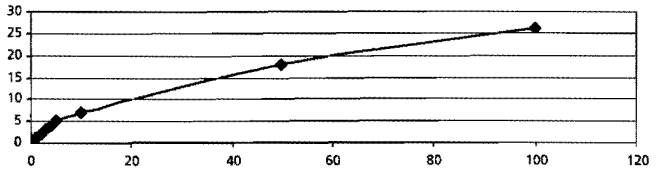
Entre esas 20 ocurrencias, hay palabras que se repiten: *ábreme* (dos veces), *la* (cuatro veces), *puerta* (tres veces).

b] Llamaremos *tipo* a cada palabra encontrada, eliminando de nuestra cuenta sus repeticiones, por lo que tendremos 14 tipos.

c] Como sabemos que el vocablo es la forma representante de un paradigma de palabras, nuestros 14 tipos se reducirán a 13 vocablos, pues *ábreme* y *abrirte* son dos formas (dos *tipos*) del vocablo *abrir*. El *vocablo* es la unidad de léxico.

Con esas distinciones, procedemos al análisis del corpus y veremos que, conforme se va analizando un texto y después un conjunto de textos, el número de ocurrencias va creciendo una a una hasta terminar su recuento; en cambio, el número de tipos va creciendo menos, porque muchos de ellos se repiten, y el número de vocablos resultante aumenta todavía menos, porque todos los tipos pertenecientes al mismo paradigma cuentan como un solo vocablo. Después de analizar, por ejemplo, 100 000 *ocurrencias* de palabras, la cantidad de *vocablos* que hayamos reconocido será mucho menor.

El corpus se *vuelve suficiente* a partir del momento en que siguen creciendo las ocurrencias y es cada vez más difícil encontrar vocablos nuevos. En ese momento hay que dejar de incluir textos en el corpus, si no quiere uno trabajar de balde. La estadística nos enseña que en casos como éstos, la relación entre número de ocurrencias de un corpus y cantidad de vocablos reconocidos en él tiende a perder su proporcionalidad. Si lo traducimos a una gráfica, en la que la cantidad de ocurrencias se representa en el eje de las abscisas, y la cantidad de vocablos en el eje de las ordenadas, lo que se produce es una "curva asintótica", que es aquella que tiende a volverse paralela al eje de las abscisas.



Así por ejemplo, en el *Corpus del español mexicano contemporáneo*, que consta de 1 891 045 ocurrencias de palabras, encontramos 64 183 tipos y cerca de 55 000 vocablos. Si comparamos este resultado con el obtenido por el corpus del *Trésor de la langue française*, de 70 millones

de ocurrencias, del que se obtuvieron 76 000 vocablos, veremos que la suficiencia de nuestro primer corpus se alcanzó comparativamente con menos datos, por lo que, desde el punto de vista de la representatividad, nuestro corpus mexicano es tan representativo como el francés, sólo que menos costoso; es decir, resultó más eficaz. Dicho de otra manera: la suficiencia de un corpus depende de su riqueza léxica, más que de su tamaño medido en ocurrencias.

Un corpus nunca podrá ser lo suficientemente grande como para poder reconocer en él todo el léxico de la lengua histórica; es sólo una muestra de ella que, sin embargo, ofrece muchas ventajas, que se irán exponiendo en este capítulo. Un corpus de una lengua muerta, del latín, del griego antiguo o del náhuatl del siglo XVI, es limitado, pues nadie creará nuevas palabras en esas lenguas. Sólo en ese caso el corpus es exhaustivo y el léxico determinado.

Más adelante consideraremos otras formas de recolección de datos léxicos.

2.2.3. La transcripción

Ya hemos insistido en el valioso papel que tiene la escritura para conservar la lengua y poder estudiarla con detenimiento. Cuando se trabaja con grabaciones del habla viva es necesario transcribirlas, lo cual impone varias dificultades de las que hay que tener conciencia. Como podemos notar, el habla está sometida a mucha variación, debido a las condiciones en que se produce: cómo suena la voz de una persona, qué tan rápido habla, cuántos fragmentos de signos omite, cuántas veces repite algo o retrocede para modificar sus oraciones, qué pronunciación tiene, cuántas oraciones quedan sin terminar, cuántas veces se equivoca en la concordancia temporal, etcétera.

Como se enseña en fonética, es posible hacer una *transcripción fina*, mediante un alfabeto fonético, que trate de recuperar los matices más sutiles de la pronunciación, o una *transcripción gruesa*, dedicada a detallar sólo los aspectos fonéticos que interesen (por ejemplo, en español, puede no prestarse mucha atención a la pronunciación de fonemas como /t/, /p/ o /k/, mientras que las vocales, los grupos consonánticos, la pronunciación de /s/, etc. suelen requerir mucha atención; en lenguas

amerindias, que tienen un rico conjunto de pares consonánticos en que hay consonantes glotalizadas, es en cambio muy importante hacer una transcripción fina, precisamente de fonemas como /t/ y /t'/, /p/ y /p'/, /k/ y /k'/). Esta clase de transcripción fonética, fina o gruesa, suele no ser necesaria cuando el objeto de estudio es el léxico, para el que no interesa su sustancia de la expresión...¡salvo! cuando esa sustancia es la que ayuda a reconocer un vocablo en una pronunciación confusa, o cuando hay que decidir de qué homófono se trata, por ejemplo, entre <casar> y <cazar> en regiones que, como Hispanoamérica, son seseantes y no distinguen los fonemas /s/ y /θ/.

Para un estudio léxico puede, en consecuencia, hacerse una transcripción regida por el sistema de escritura de la lengua en estudio, pero en ese caso pueden aparecer problemas ortográficos como los siguientes: ¿hay que transcribir <hiede> o <jiede> cuando escuchamos [hjéde]?, ¿qué forma escoger <abalanzarse> o <avalanzarse> cuando escuchamos [abalanθarse] o [abalansarse]?, cuando oímos [witlacoche], por qué forma ortográfica optamos, ¿por la etimológica <cuitlacoche>, por otra muy común como <huitlacoche> o incluso por la popular <güitlacoche>? Tomar estas decisiones es determinante para un estudio léxico cuidadoso, pues puede tener efecto sobre los resultados ortográficos, morfológicos o lexicográficos que queramos alcanzar. De ahí la necesidad de definir los objetivos del estudio antes de comenzar a hacer la transcripción y de especificar los criterios que han de seguirse en ella.

Como todos tenemos una lengua materna por la que sentimos lealtad, y en nuestra educación se ha impuesto una concepción ortográfica tradicional poco analizada, que distingue lo correcto de lo incorrecto, podemos caer inadvertidamente en la imposición de una ortografía determinada a palabras que desconocemos o que socialmente están mal vistas. Si lo hacemos, podemos dañar la calidad de nuestros datos léxicos. Por ejemplo: sabemos que todavía tienen uso formas antiguas del español en zonas rurales, como *naiden* en vez de *nadie*, *ansina* en vez de *así*, *truje* en vez de *traje*, etc. La existencia de estas palabras es un dato muy importante, tanto para la memoria de la lengua histórica como para la caracterización de un habla rural; si al transcribirlas les imponemos las formas más modernas, dañamos nuestra investigación. Lo mismo podemos decir en relación con las diferencias dialectales: para un náhuatl de Texcoco o de Cholula, muchos vocablos del náhuatl de Guerrero o de Veracruz discreparán notablemente de los suyos; si al recoger esos datos les impone su propio conocimiento de la lengua, los resultados que obtenga quedarán irremediablemente falseados. Los mismos errores inducen falsas suposiciones etimológicas; por ejemplo, cuando en México se oye decir que una persona es [abusada], que quiere decir que es lista, inteligente, capaz, una falsa etimología propone que el vocablo sea "aguzado"; si se transcribe con esta falsedad, el resultado falla.

No es posible dar reglas mecánicas para conseguir una buena transcripción, pero sí se puede recomendar que, ante todo, definamos con la mayor claridad posible la clase de estudio que queremos hacer, para que dependan de ella nuestros criterios de transcripción, y después, que hagamos una constante crítica de nuestras propias ideas normativas acerca de la lengua, para impedir que se filtren en nuestra transcripción. Quien esté interesado en profundizar sus conocimientos acerca de problemas de transcripción, podrá leer el libro de Celia Zamudio, incluido en la bibliografía del capítulo quinto.

2.3. Corpus de estudio del léxico de grupos sociales

Hemos venido distinguiendo entre el estudio del léxico de la lengua histórica, el de cada persona y el que genéricamente hemos llamado de "grupos sociales". Estos grupos pueden ser hablantes de una variedad dialectal, por ejemplo, los habitantes de Tlacotalpan, Veracruz, los del barrio de La Boca en Buenos Aires, los de la Mixteca alta, etc.; también pueden ser los jóvenes de una región de una ciudad o los de cierta clase socio-económica; igualmente los albañiles, los médicos o los delincuentes que hablan una germanía o caló. En estos casos, el corpus se define a base de la caracterización de los hablantes, no de los textos, como cuando estudiamos la lengua histórica. Se podría pensar que, en consecuencia, aquí sí conviene aplicar cálculos de representatividad como los de las muestras electorales, pues la "población total" la conocemos, ya que no es de vocablos sino de personas. En todo caso, siempre hay que tratar de que la muestra sea reducida, cuidando la aleatoriedad de la selección de hablantes para reducir al máximo el riesgo de que sea una muestra sesgada.

La muestra debe estar formada al menos por tres representantes de cada variable socio-económica que pueda influir en el vocabulario: de

grupos de edad, de escolaridad, de sexo, de ocupación, si mira televisión o escucha radio, si lee prensa y libros, etc. para asegurar la validez social de los datos; es posible combinar a los informantes, de modo que cada uno de ellos se tome en cuenta para diferentes variables; así por ejemplo, un hombre, de edad madura, con estudios elementales y carpintero de oficio, vale como informante para esas cuatro variables. Recordemos también que las grabaciones que se hagan de diálogo espontáneo ofrecen datos del vocabulario activo de cada persona, no de su vocabulario pasivo, de modo que los resultados que se obtengan serán *caracterizadores* del grupo, pero no exhaustivos ni responderán a la totalidad del léxico que conozca ese grupo.

2.3.1. La grabación

No se debe olvidar, y menos cuando lo que sustenta el trabajo es una concepción pragmática de la lingüística, que las personas hablan para actuar. No se puede pedir a nadie, delante de un micrófono "que diga algo"; nuestra presencia nos vuelve sus interlocutores y debemos corresponder a las condiciones de diálogo que se producen en todo encuentro. Cuentan, por lo tanto, el modo en que nos presentemos, el vestido o las joyas que tengamos puestas y nuestra propia manera de hablar como *síntomas* nuestros, que influyen y pueden sesgar las respuestas de nuestros informantes; cuenta también la necesaria cortesía entre ellos y nosotros, y los límites de corrección social que existan en la sociedad estudiada.

Así por ejemplo, un hombre no puede proponer a una mujer un diálogo sobre temas íntimos de las mujeres, o una mujer no puede entrar a una cantina a buscar informantes hombres, en las condiciones sociales de un país como México; no puede uno imponer una entrevista conducida en español para obtener de ella vocabulario de una lengua amerindia; no puede interrumpir tareas de los informantes, ni tratar de que hablen de actividades en las que acostumbran estar callados.

Una conversación o un interrogatorio con cuestionario debe hacerse después de haber logrado una aceptación del informante, y siempre tratando de que la conversación sea espontánea, flexible e interesante. Sólo cuando se cumplen estas condiciones se puede proceder a hacer una grabación.

2.3.2. Los textos

Es difícil obtener textos de una persona que no tenga una ocupación que la obligue a escribir; los únicos que a veces es posible conseguir son cartas personales, pero en ese caso debe uno cuidar mucho la intimidad de emisor y destinatario y obtener un permiso explícito de ellos para utilizar las cartas como material de estudio.

2.3.3. Los cuestionarios

Acostumbrados por el método dialectológico, los cuestionarios suelen estar compuestos por una serie de "conceptos" (que, estrictamente hablando, no lo son, sino vocablos de una lengua utilizada como base de la observación —véase el capítulo 9, § 4—) cuyos nombres se busca encontrar en una comunidad lingüística determinada. Son, en consecuencia, cuestionarios de vocablos nominales, lo cual reduce su utilidad para un estudio del léxico. Además, aunque se consideren "conceptos" son en realidad vocablos de la lengua de su autor, por lo que no dejan de estar cargados del punto de vista correspondiente a su cultura. Esa clase de cuestionarios debe manejarse siempre, por eso, con extremo cuidado y conocimiento previo de lo que se está preguntando y del modo de preguntarlo.

Otra forma de obtener designaciones de objetos es proponer un tema de conversación e impulsar a su informante a que hable de él. Así por ejemplo, si se quiere reunir vocabulario relativo a la siembra de maíz, se propone hablar de ello y se graba todo el relato que haga el informante al respecto, incluso participando uno mismo en esa tarea.

2.4. Corpus cualitativos

Hasta aquí hemos venido hablando de corpus que ofrezcan una base documental representativa de una lengua histórica o de una variedad regional o social. En todos estos casos, el estudio lexicológico que hagamos nos permitirá llegar a conclusiones legítimas y verificables acerca de la realidad de la parcela léxica que queremos estudiar. Por ejemplo, el

Corpus del español mexicano contemporáneo nos permite afirmar que los datos que obtuvimos forman parte realmente del español de México, en tanto que no podemos afirmar que sean un documento del léxico total del español de México. Si construimos un corpus de cartas personales de habitantes de la ciudad de Jerez, Zacatecas, podemos afirmar que los vocablos obtenidos forman parte del léxico de ese grupo social, pero no podemos afirmar que ese sea todo el léxico del grupo; si tomamos como corpus el conjunto de cartas escritas en maya, correspondientes al período histórico de la guerra de las castas en Yucatán, el vocabulario que obtengamos es una muestra del vocabulario maya de esa época, pero no todo su vocabulario.

2.4.1. Corpus de finalidad exhaustiva

En todos estos casos, el léxico nos interesa por su representatividad de un estado de una lengua; sin embargo, también nos puede interesar construir un corpus que trate de ser exhaustivo, en relación con una totalidad de documentos de una lengua histórica o de un estadio de su evolución. Por ejemplo, la Academia Española, que tiene ya casi 300 años de haberse fundado y que desde entonces ha venido reuniendo un valioso acervo de datos, ha formado un gran corpus de la lengua histórica española con vocación de exhaustividad (aunque no lo vaya a lograr, por las razones antes expuestas); un corpus de esta clase no se plantea su representatividad, sino sólo su exhaustividad, por lo que vale lo mismo registrar en él un vocablo usado una sola vez (un aparente *hapax*), que uno que se repite miles de veces; un vocablo utilizado por toda una comunidad, que uno utilizado por un solo autor; un vocablo del siglo XIII, que uno del XXI.

El objetivo del corpus es más filológico que lingüístico; los vocablos encontrados valen por sí mismos, en su propia peculiaridad, y no como representantes del vocabulario de una comunidad. Esa clase de corpus la llamaremos cualitativa. El *Corpus diacrónico del español* (Corde) y el *Corpus de referencia del español actual* (CREA) de la Academia Española son dos buenos ejemplos de corpus con valor cualitativo. En el capítulo siguiente trataremos sus ventajas y desventajas desde el punto de vista de su tamaño y sus características.

El estudio de un solo vocablo o de una pequeña cantidad de ellos es tan legítimo como el estudio de una gran masa de vocablos pertenecientes a una comunidad lingüística. Acerca de esta clase de estudios hablaremos en el capítulo 9.

2.4.2. Las "concordancias"

Ciertos corpus, en los cuales el objetivo central no es el estudio del vocabulario, pero que se prestan muy bien para hacerlo, son los que se conocen como "concordancias". Desde hace mucho tiempo se habían venido preparando índices de vocablos contenidos en la obra de diferentes autores, con muy distintas finalidades; por ejemplo, el jesuita italiano Roberto Busa preparó un *Index thomisticus* dedicado a la obra de Santo Tomás de Aquino, evidentemente, con el propósito de facilitar el estudio teológico y filosófico de este padre de la Iglesia católica.

Esta clase de índices consta fundamentalmente de una lista de vocablos o de expresiones utilizados en una obra, seguidos de referencias a los lugares en que aparecen y, además, de una o más líneas de texto, que ofrecen el contexto que facilita su interpretación. Desde que se introdujo a la lingüística el uso de la computadora electrónica —precisamente Busa fue uno de sus iniciadores—, la práctica de producir concordancias de la obra de diversos autores se volvió común. Hoy en día la obra de muchos escritores ha sido objeto de concordancias; de Cervantes, del Arcipreste de Hita, de García Lorca, etc. Para estudios de vocabulario estos trabajos son muy útiles, pero además, la práctica se ha extendido a la lingüística; los tres corpus citados antes, de la Academia Española y del español de México producen concordancias, que resultan muy útiles para llevar a cabo estudios léxicos.

3. CONCLUSIÓN

Como se ha mostrado, la lingüística dispone de varios métodos de recolección de datos léxicos, cuya selección depende de los objetivos del estudio que se desee realizar. En este capítulo nos hemos ocupado de los conceptos básicos para definir la conveniencia y las características de los

métodos de recolección de los datos léxicos; en el siguiente discutiremos el carácter cuantitativo del vocabulario fundamental, del vocabulario disponible y de un método más para reunir datos valiosos para el estudio del significado de los vocablos.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

vocabulario fundamental; vocabulario activo; vocabulario pasivo; vocabulario disponible; léxico común; *hapax legomenon*; representatividad; ocurrencia; tipo; transcripción; corpus; concordancia.

OBRAS CONSULTADAS

LARA, Luis Fernando, Isabel GARCÍA HIDALGO y Roberto HAM, *Investigaciones lingüísticas en lexicografía*, México, El Colegio de México, 1979.

LOPE BLANCH, Juan, *Cuestionario para la delimitación de zonas dialectales de México*, México, El Colegio de México, 1970.

MULLER, Charles, "Le mot, unité de texte et unité de lexique en statistique lexicologique", *Travaux de Linguistique et Littérature*, 1(1963), pp. 155-173.

Corpus:

Corpus de referencia del español actual, Real Academia Española. Se puede consultar en www.rae.es

Corpus diacrónico del español, Real Academia Española. Se puede consultar en www.rae.es

Corpus del español mexicano contemporáneo, México, El Colegio de México, 1974. Puede consultarse mediante solicitud expresa al Diccionario del Español de México, El Colegio de México.

8
ESTUDIOS CUANTITATIVOS DEL LÉXICO

INTRODUCCIÓN

La lingüística ha sido desde sus orígenes una ciencia predominantemente cualitativa y ha considerado la cuantificación en general sólo como un instrumento valioso de recolección de materiales, sin poderle reconocer o asignar en sus especulaciones acerca de la naturaleza de las lenguas un papel más importante, que revele una probable naturaleza cuantitativa de los sistemas lingüísticos. En este capítulo nos ocuparemos de algunos métodos de estudio cuantitativo del léxico siguiendo los cauces establecidos por la lingüística, pero tratando de llamar la atención hacia el papel que puede llegar a tener la cuantificación en la comprensión general de la naturaleza del léxico.

1. CONCEPTOS BÁSICOS

En el capítulo anterior propusimos que un corpus debe ser suficientemente grande y rico en vocablos diferentes, para que pueda darnos resultados que sean representativos del léxico de la lengua estudiada. Reunir un corpus muy grande de una lengua como la española se ha vuelto cada día más sencillo, pues basta pasar los textos seleccionados por un escáner e introducirlos en una base de datos electrónica, para que el corpus vaya creciendo hasta quedar constituido por muchos millones de ocurrencias (no es nada sencillo reunir un corpus de alguna lengua amerindia debido a la falta de textos suficientes y de grabaciones adecuadas, que caracteriza, desgraciadamente, la investigación de esas lenguas).

Ya vimos que no es necesario un corpus de gran tamaño para alcanzar un registro suficiente del léxico de una lengua histórica en un momento dado de su historia. Ahora podemos agregar que ese corpus

debe estar compuesto por textos de muchos autores diferentes y de muchos tipos o géneros (novelas, teatro, cartas personales, diarios, ensayos, recetarios de cocina, novelas del corazón, textos de física, de astronomía, de biología, de ingeniería, de religión, etc.) pues de no ser así, se producirían sesgos importantes en la aparición de los vocablos en el corpus, que dañarían irremediablemente la representatividad del estudio.

Tomemos por caso lo que sucedería si decidiéramos tomar como corpus sólo los textos periodísticos de José Alvarado, un autor muerto hace cerca de 20 años, que colaboraba en varios diarios y revistas mexicanos; seguramente podríamos reunir varios cientos de miles de ocurrencias a partir de sus textos, correspondientes al español mexicano culto de finales del siglo XX, por lo que tendríamos un corpus suficientemente grande. Lo que no podríamos prever es que ese autor escribía con un curioso prurito: nunca introducir oraciones subordinadas con la conjunción *que*. El resultado que obtuviéramos, sólo por ese faltante —en comparación con la realidad del español— sería falso.

Supongamos que también incluyéramos en el corpus las obras del escritor mexicano Fernando del Paso; en particular su novela *José Trigo*. En esa novela, Del Paso hace gala de una enorme riqueza léxica, formada por vocablos que, aunque incluidos en varios diccionarios, no se llegan a documentar en otros textos del español contemporáneo. El resultado sería la aparición anómala, en comparación con la realidad de la lengua, de muchos vocablos, existentes y registrados en los diccionarios, pero de un uso muy restringido en la vida social. Es decir que el estilo, las preferencias y aun los pruritos extraños de cada autor —a los que sin duda tiene derecho— tienden a sesgar los resultados cuantitativos de un corpus en una sola dirección, lo que les resta representatividad.

Supongamos ahora que nuestro corpus se restringiera a obras dedicadas a la cardiología. La mayor parte del vocabulario encontrado en ellas correspondería seguramente al estado de la lengua considerado, pero la frecuencia de aparición de voces como *corazón*, *vena*, *sangre*, *infarto*, etc. podría ser tan grande, que también falsificaría nuestros resultados. Por eso es por lo que un corpus debe construirse con la mayor variedad posible de temas y géneros del discurso y del texto.

Hagamos el siguiente experimento sencillo: supongamos que en un corpus hemos reunido textos de cinco géneros diferentes y de muchos autores distintos. Hemos obtenido los siguientes resultados en términos de ocurrencias:

	cartas	*teatro*	*ciencia*	*periodismo*	*religión*
casa	10	6	2	4	4
correr	3	4	0	5	1
epilepsia	0	0	15	3	0
escindir	0	1	5	2	4
filólogo	0	0	2	1	3
pecado	3	2	0	1	15
tomar	10	8	6	8	9

1.1. Frecuencia y dispersión

Llamaremos *frecuencia de un vocablo* al número de ocurrencias que registramos de él en un corpus; la representaremos con la letra F.

Llamaremos *dispersión de un vocablo* a la manera en que se reparte su aparición entre todos los géneros del corpus; la representaremos con la letra D.

Los vocablos de nuestro ejemplo tendrían las siguientes características cuantitativas: casa (F = 26, disperso en los 5 géneros), correr (F = 13, en 4), epilepsia (F = 18, en 2), escindir (F = 12, en 4), filólogo (F = 6, en 3), pecado (F = 21, en 4) y tomar (F = 41, en 5). Nuestra percepción subjetiva del uso de la lengua nos dirá que *casa* y *tomar* han de ser vocablos completamente usuales en español, y sus frecuencias, así como su dispersión, nos lo confirman; de *epilepsia*, en cambio, pensaríamos que es menos usual que los anteriores; nuestros resultados indicarían que, aunque alcanza una frecuencia comparativamente alta en el género de ciencia, se utiliza menos en el resto de nuestros textos; *pecado* ha resultado

más frecuente que *correr* —lo que nos sorprende (y no porque se peque menos de lo que se corre)— y *escindir*; *filólogo* ha resultado con poca frecuencia y menos dispersa. Si quisiéramos concretarnos a sumar las ocurrencias de los vocablos en un corpus, los resultados de frecuencia producirían el siguiente orden: *tomar*, *casa*, *pecado*, *epilepsia*, *correr*, *escindir* y *filólogo*.

Si quisiéramos corregir esos resultados incorporándoles una medida de su dispersión, tendríamos que buscar una fórmula que nos lo permitiera. La manera simple de poner en relación los resultados de frecuencia y dispersión de los vocablos es multiplicar la frecuencia por el número de géneros en que aparecen, como una medida de su dispersión, lo que nos daría el *uso* (U) de cada vocablo (adoptando un término técnico utilizado, aunque de manera más elaborada, por Alphonse Juilland y Emilio Chang Rodríguez, quienes fueron de los primeros en introducir cálculos cuantitativos en la investigación lingüística en su *A Frequency Dictionary of Spanish Words*).

$$U = F \times D$$

Si seguimos esta fórmula, el orden de los vocablos encontrados sería el siguiente: *tomar* (U = 205), *casa* (U = 130), *pecado* (U = 84), *correr* (U = 52), *escindir* (U = 48), *epilepsia* (U = 36) y *filólogo* (U = 18). Comparando este orden con el de la frecuencia simple, vemos que *epilepsia* bajó y, en cambio, *correr* y *escindir* se acercaron a las tres primeras, como podría corresponder a nuestra percepción subjetiva del mayor o menor uso de las palabras.

Lo que se busca con estos cálculos es encontrar un procedimiento de análisis cuantitativo que mida la realidad del uso de los vocablos de una lengua en un estadio histórico determinado. La frecuencia es el primer dato que se puede obtener, pero es igualmente importante medir de manera adecuada la forma en que se reparte un vocablo en diferentes géneros discursivos y de texto, es decir, su dispersión.

No abundaremos aquí en las diferentes medidas que se han propuesto de la dispersión y el uso de los vocablos, porque requieren unos conocimientos básicos de estadística que no podemos ofrecer en este *Curso*. Quien esté interesado en ellas puede consultar las obras que se listan al final de este capítulo.

1.2. Frecuencias absoluta y relativa

La totalidad de las ocurrencias de un vocablo en un corpus nos ofrece su *frecuencia absoluta*. Puede ser muy útil considerar la frecuencia que tenga un vocablo en cada género textual o discursivo. En el ejemplo anterior, las *frecuencias relativas* a cada género son las que se listan bajo sus respectivos encabezados.

1.3. Porcentajes de frecuencias

Una vez que uno conoce las frecuencias relativas de un vocablo en cierto género, puede querer saber qué tan importantes son sus vocablos en comparación con los demás. En tal caso, mediante una sencilla regla de tres, se obtiene el porcentaje de frecuencias del vocablo *x* respecto del total de ocurrencias de vocablos en el género *Y*.

1.4. Ponderación de géneros discursivos o textuales

Cuando uno maneja un corpus léxico de varias decenas de miles de vocablos y varios millones de ocurrencias, las comparaciones no son tan sencillas como las de nuestro ejemplo. Juilland y Chang Rodríguez construyeron un corpus de 500 000 ocurrencias, divididas en cinco géneros de 100 000 ocurrencias cada uno. Como todos los géneros eran del mismo tamaño, los porcentajes de frecuencias relativas en cada género eran comparables de género a género.

Ahora supongamos que nuestro corpus, además de estar dividido en muchos géneros diferentes, con textos de muchos autores o de muchas procedencias, también trata de asignar importancia relativa a cada género. Por ejemplo, en la vida social se lee más la prensa que el tratado científico; las conversaciones juveniles son menos abundantes que las conversaciones entre señoras; las obras literarias se leen más de lo que se escuchan los sermones y las prédicas religiosas, etc. Los géneros del corpus adquieren así diferentes ponderaciones, que se traducen en cantidad de textos incluidos en cada uno. Así por ejemplo, en el *Corpus del español mexicano contemporáneo* incluimos 150 textos literarios fren-

te a sólo 55 de conversaciones coloquiales; 180 textos científicos frente a 30 de jergas juveniles y de la delincuencia.

El efecto de introducir ponderaciones de los géneros es que se vuelven incomparables los datos de frecuencia y dispersión entre los géneros del corpus, pues no es lo mismo que un vocablo tenga alta frecuencia en un género compuesto por 150 textos, a que la tenga en otro de 30 textos. La ponderación de los géneros interfiere en el cálculo de la dispersión, tal como lo explicamos antes.

Si uno quiere introducir ponderaciones de esta clase, los cálculos cuantitativos se vuelven complicados y hace falta tomar un curso de estadística lingüística o auxiliarse con los conocimientos especializados de un estadígrafo. Por tal motivo no continuaremos explicando aquí cómo se llevan a cabo esos cálculos. Lo que es importante es saber que todo estudio cuantitativo del léxico requiere conocer bien la frecuencia absoluta de cada vocablo, sus frecuencias relativas junto con sus porcentajes, la influencia de las características de construcción del corpus de datos y el cálculo de la dispersión de cada vocablo en el corpus.

Veamos los resultados cuantitativos reales que obtuvieron los siete vocablos del ejemplo en nuestro *Corpus del español mexicano contemporáneo*. Nos limitaremos a considerar solamente dos tipos de datos: el de frecuencia absoluta del vocablo en todo el corpus y el de dispersión, que en este caso se ha sintetizado en un "índice normalizado de dispersión", simbolizado con la letra C, y que es resultado de una complicada fórmula que toma en cuenta el tamaño de todo el corpus, los tamaños de cada género, las frecuencias absolutas de los vocablos, sus frecuencias relativas en cada género y el modo en que se reparte su uso entre los géneros, con la ventaja de que coloca la dispersión en una escala de cero a uno: mientras mejor repartido está el vocablo en el corpus, más se acerca al valor de uno.

	F	*C*
tomar	1 606	0.98
correr	459	0.84
casa	1 820	0.83
pecado	77	0.48
escindir	2	0.31
filólogo	3	0.31
epilepsia	6	0.28

Como se ve, la frecuencia absoluta se corrige mediante el cálculo de su dispersión y no resultan tan determinantes las diferencias de frecuencia. Si *correr* tiene cuatro veces menos ocurrencias que *casa*, su dispersión es casi la misma que la de ésta; si *epilepsia* tiene mayor frecuencia absoluta que *escindir* y *filólogo*, la ponderación de los géneros del corpus permite tomar en cuenta que los dos últimos tienen una dispersión ligeramente mejor que el primero. A la vez, *tomar* se manifiesta como un vocablo muy frecuente y casi perfectamente disperso en todo el corpus.

Con esos conocimientos básicos, pasemos ahora a aplicarlos a los métodos cuantitativos más usuales de estudio del léxico.

2. EL CÁLCULO DEL VOCABULARIO FUNDAMENTAL

En el capítulo anterior explicamos qué se entiende por *vocabulario fundamental* de una lengua histórica. Ahora pasaremos a analizar cómo se determina.

Un vocabulario fundamental sólo puede ser resultado de una amplia exploración cuantitativa de un corpus de datos léxicos. Aun cuando el vocabulario fundamental, por serlo, tiende a aparecer en cualquier corpus, incluso si es reducido, es necesario asegurarse antes de que el corpus contenga una amplia variedad de textos, que permitan registrar ese vocabulario de una manera segura y cuantitativamente medible, si no quiere uno fallar tanto en la detección de ese vocabulario como en su necesaria representatividad social.

No hay un criterio único para determinar el vocabulario fundamental, ni una sola fórmula para calcularlo. Víctor García Hoz, que fue el primero en llevar a cabo estos cálculos sobre la lengua española, consideraba que el vocabulario fundamental es el conjunto de vocablos de una lengua histórica cuyas frecuencias aparecen prácticamente distribuidas en partes iguales en todos los géneros de un corpus. Para encontrarlo, construyó un corpus de 400 000 ocurrencias, dividido en cuatro géneros del mismo tamaño (100 000 ocurrencias cada uno): cartas, periódicos, documentos oficiales y obras literarias. Obtuvo de él 12 913 vocablos; 1 971 de ellos con frecuencia mayor o igual de 40, al que denomina "vocabulario común" (o sea, es la intersección de los cuatro conjuntos de vocablos con frecuencia mayor o igual a 10). De ese vocabulario común seleccionó aquellos vocablos cuya frecuencia absoluta fuera exclusivamente de 40 ocurrencias, correspondiente a 10 ocurrencias en cada género, perfectamente repartidas. Su resultado fue un *vocabulario fundamental* de 208 vocablos.

Si uno revisa este vocabulario comprueba que no incluye los pronombres personales *tú*, *ella*, *nosotros*, *vosotros* y *ellos*; tampoco las preposiciones *ante*, *bajo*, *contra*, *hacia* y *tras*. Estos vocablos al menos, que son imprescindibles para construir oraciones en español, revelan cierta deficiencia del corpus empleado, y además le restan utilidad para cualquier aplicación práctica en la enseñanza de la lengua. El "vocabulario común" de García Hoz, constituido por todos los vocablos con frecuencia absoluta igual o mayor de 40 ocurrencias, es de mayor utilidad.

Georges Gougenheim y sus colegas del Centre de Recherche et d'Étude pour la Diffusion du Français tomaron como criterio central para definir su vocabulario fundamental la frecuencia absoluta de los vocablos en un corpus formado por 163 textos transcritos de conversaciones de personas de diferentes regiones francesas, distintos oficios y profesiones y temas variados (aunque desgraciadamente poco fragmentado en géneros que permitieran un análisis más cuidadoso y adecuado de la dispersión), que dieron por resultado 312 135 ocurrencias, correspondientes a 7 995 vocablos. Sobre esa base añadieron posteriormente sus resultados de un estudio de vocabulario disponible (al que nos referiremos más tarde) y una serie de inclusiones y exclusiones más, al léxico que habría de servir para estructurar y orientar la enseñanza de la lengua francesa. Arbitrariamente consideraron que todo vocablo con $F \geq 20$ formaría parte de su vocabulario fundamental. El vocabulario fundamental del francés quedó compuesto por 1 475 vocablos.

Para determinar el vocabulario fundamental del español de México seguimos un procedimiento diferente: tomamos como límite los vocablos que constituyen el 75% del total de las ocurrencias en el corpus. En un estudio por cuartiles (acumulaciones del 25% de las ocurrencias), encontramos que el primer cuartil lo forman las ocurrencias de sólo 9 vocablos; el segundo (50%), de 74 vocablos; el tercero (75%), de 1 131 vocablos. La acumulación por cuartiles del vocabulario de la lengua culta fue, correspondientemente, de 9, de 100 y de 1 451 vocablos.

Como se puede ver, no hay un criterio fijo para determinar un vocabulario fundamental. Todos los estudios al respecto se hacen al tanteo, tomando en cuenta el objetivo que se persigue, por ejemplo, en la enseñanza de lenguas. Sin embargo, el hecho de que diferentes estudios cuantitativos del léxico muestren cómo el número de vocablos de una lengua que se utilizan frecuentemente y en toda clase de acciones verbales oscila entre los 1 000 y los 2 000, es un indicador valioso de las características del uso de las lenguas y del modo en que opera la memoria humana.

3. EL ESTUDIO DEL VOCABULARIO DISPONIBLE

El mismo grupo de investigación de Georges Gougenheim citado antes puso en práctica la investigación del vocabulario disponible. Como dijimos en el capítulo anterior, este vocabulario consiste en aquel conjunto de vocablos que cada persona puede utilizar activamente en cualquier acción verbal, bajo el estímulo de sus necesidades específicas de comunicación. A diferencia del vocabulario fundamental, que es aproximadamente el mismo para todos los hablantes de una lengua histórica, el vocabulario disponible varía según los conocimientos, las actividades y la educación de cada persona. Por el hecho de que tal conjunto de vocablos no se utiliza necesariamente en todas las ocasiones, no tiene una alta frecuencia de aparición, motivo por el cual su registro en un corpus tiende a disminuir, junto con su capacidad de dispersión. Gougenheim y sus colegas utilizaron su investigación del vocabulario disponible en francés para mejorar los resultados obtenidos en el estudio del vocabulario fundamental y poder ofrecer un léxico más completo a los aprendices de la lengua francesa.

Decíamos en el capítulo anterior que el vocabulario disponible es una manifestación asociativa de la memoria de las personas. Si se propone un tema determinado, llamado *centro de interés* en los estudios de esta clase, toda persona es capaz de recordar y expresar palabras que considera forman parte de ese centro. Por ejemplo, si se propone como centro de interés "animales salvajes", una persona puede emitir una lista como la siguiente: *león*, *tigre*, *oso*, *jabalí*, *jaguar*, etc.; si se propone como tema "las partes del cuerpo", enunciará *cabeza*, *brazo*, *mano*, *pierna*, *pie*, etc. Cada persona *dispone* de vocabularios correspondientes a tantos centros de interés como sea diversa su actividad de conocimiento y su educación, por lo que es probable que su vocabulario disponible esté compuesto por varios miles de vocablos.

El modo de investigar esta *disponibilidad léxica* suele consistir en pedir a cierto número de informantes que, en un corto período (dos a tres

minutos), emitan o escriban todas las palabras que les vengan a la memoria, asociadas con un centro de interés determinado. Que el período dedicado a la emisión o a la confección de una lista de palabras sea corto, es importante en cuanto garantiza la espontaneidad del recuerdo, pues dificulta la reflexión.

Los investigadores franceses mencionados prepararon una lista de centros de interés formada por los siguientes: partes del cuerpo, vestido, casa, muebles de la casa, alimentos y bebidas de las comidas diarias, objetos colocados en la mesa cuando se come, la cocina, sus muebles y utensilios, la escuela, sus muebles y materiales, calefacción y limpieza, la ciudad, el pueblo, medios de transporte, trabajos del campo y del jardín, animales, juegos y distracciones, oficios. Los investigadores que se han dedicado a estos estudios en el ámbito de la lengua española conservaron los mismos centros de interés franceses, con las siguientes modificaciones: electricidad y aire acondicionado, la naturaleza, diversiones y deportes, según lo relata Juan López Chávez en su libro *¿Qué te viene a la memoria? La disponibilidad léxica: teoría, métodos y aplicaciones.*

Aunque esa haya sido la costumbre hasta ahora, todo tema puede ser centro de interés, y su selección depende del objetivo que se quiere alcanzar. Si, por ejemplo, se quiere saber qué vocabulario comprenden y quizá manejan los estudiantes de preparatoria en cuanto al centro de interés de biología, la encuesta se organiza a partir de ese centro.

De los resultados obtenidos por los grupos francés y mexicano, resulta notable que los miembros de cada grupo de informantes, clasificados por edades, por oficios, por nivel de escolaridad, por región, etc. en un estadio contemporáneo de la lengua histórica, *tienden a recordar los mismos vocablos y aproximadamente en el mismo orden.* Es decir que, aun cuando la asociación mental es un fenómeno individual, los miembros de cada grupo de informantes coinciden en el mismo tipo de asociaciones, lo cual es un importante indicio de la socialidad del léxico en un momento determinado de la lengua histórica.

Siendo así, el siguiente paso es la búsqueda de instrumentos de medición de la disponibilidad léxica, que puedan proporcionar una ponderación objetiva y válida de ese léxico, orientado a su uso en métodos de enseñanza de lenguas, de examen de riqueza léxica en ciertas personas, así como para cierto tipo de cuestionarios útiles para el diagnóstico neurológico.

3.1. Las medidas de la disponibilidad

De un cuestionario de disponibilidad léxica se obtienen los vocablos listados en *orden* de recuerdo; a tal orden se le busca asignar un *grado de disponibilidad*, según el número de respuestas coincidentes de cada vocablo (frecuencia). Por ejemplo: en el estudio francés, las palabras *oeil* 'ojo', *oreille* 'oreja', *nez* 'nariz', *bras* 'brazo' y *jambes* 'piernas' aparecieron como las cinco primeras de una lista de 20 (orden), sólo que *oeil*, la primera de ellas en tres de los cuatro departamentos (dispersión), apareció en cinco cuestionarios, mientras que *jambes*, en quinto lugar, tuvo una frecuencia de 20 (grado de disponibilidad). El estudio del grupo de Gougenheim optó por recoger los vocablos con mayor grado de disponibilidad para completar la elaboración del "francés fundamental".

Los estudios de la misma clase llevados a cabo por Juan López Chávez en México muestran diferentes resultados en el centro de interés de las partes del cuerpo, pues los cinco primeros vocablos fueron *cabeza, ojo, brazo, mano* y *nariz*. López Chávez se ha preocupado por llegar a establecer una medida de la disponibilidad léxica que englobe todos los factores citados, lo cual lo ha conducido a fórmulas estadísticas que dan por resultado un "índice de disponibilidad léxica" de cada vocablo en relación con cada grupo de informantes e incluso una medida de la "competencia léxica" de cada persona, establecida en relación con los resultados obtenidos en un grupo de informantes. De nuevo, como se trata de fórmulas que requieren conocimientos estadísticos, no las trataremos aquí, sino que remitimos al estudio citado en la bibliografía.

3.2. Disponibilidad y competencia léxica

Por las características de la encuesta de disponibilidad, el léxico que se obtiene es generalmente de sustantivos y a veces de adjetivos, lo cual restringe el valor de estas investigaciones para conocer el léxico de una persona o de un grupo de personas.

Sea como sea, esta clase de investigaciones, además de su valor para la elaboración de métodos de enseñanza de segundas lenguas, se puede utilizar también para diagnosticar los conocimientos léxicos adquiridos en la escuela y hacer comparaciones entre grupos escolares, grupos

de edad, escuelas y regiones dialectales de una lengua histórica. Pero además, ofrecen datos objetivos para facilitar el análisis cualitativo del léxico, como se explicará en el siguiente capítulo. Cabe suponer también que la manifestación de la disponibilidad léxica en cada persona refleje fenómenos neurológicos importantes, en particular de la manera en que reside y actúa la lengua en el cerebro. De ahí el interés de esta clase de investigaciones para la psicología y la neurología.

4. EL ANÁLISIS EN CONGLOMERADOS

Como dijimos al comienzo de este capítulo, los estudios cuantitativos del léxico se entienden sobre todo como instrumentos de investigación orientados a ofrecer una base objetiva y empírica al estudio del vocabulario, pues de otra manera, la recolección de vocablos necesaria para el estudio cualitativo quedaría sujeta al gusto de cada investigador, a sus preferencias o incluso a sus actitudes normativas (de las que se hablará en el capítulo 10).

Todo estudio cuantitativo del léxico, sin embargo, es un cálculo de la cantidad de vocablos que manejan o conocen sus hablantes; los datos de frecuencia y dispersión de cada vocablo, junto con los conjuntos de vocablos que se pueden formar a base de las mismas frecuencias y de los mismos rangos de dispersión, muestran interesantes aspectos de la manera en que se agrupa el vocabulario en un estadio determinado de una lengua histórica, en un grupo social o en cierto tipo de discursos o de textos. De allí su valor para formarse una idea de las características del léxico en la memoria de las personas, así como de sus propiedades cuantitativas en el sistema de la lengua.

Ahora veamos un interesante enfoque cuantitativo del léxico, que ofrece nuevas perspectivas para la formación de una base objetiva de datos, útil para el posterior análisis cualitativo, pero cuyo verdadero objeto es contribuir a medir las relaciones de significado que se forman entre las palabras de un texto.

El lingüista alemán Burghard Rieger utilizó un corpus de periódicos (*Die Welt* y *Neues Deutschland*) de las antiguas Alemania Occidental y Alemania Oriental (esta última, la República Democrática Alemana, ya disuelta, cuyos estados hoy forman parte de la primera, la República Federal de Alemania), con el objetivo de estudiar el vocabulario contenido en esos textos con base en ideas procedentes de la lógica difusa (*fuzzy logics*), y ofrecer una medida y una representación de la mayor o menor cercanía semántica entre sus miembros, lo que constituye una versión diferente de los estudios de vocabulario disponible que, sin embargo, permite sacar conclusiones a propósito de la manera en que un "centro de interés" o el sentido de un discurso o de un texto son los que determinan las relaciones de significado entre las palabras que los componen.

A diferencia de la estructura de corpus que se ha venido preconizando aquí, compuesta por textos de autores diferentes y de procedencias muy variadas, el corpus utilizado por Rieger es una colección homogénea de textos periodísticos completos; eso le permite tomar el sentido completo de los textos como hilo conductor de la medición del significado de los vocablos que los componen y aceptar la influencia (el sesgo) de los temas, del estilo de los autores y quizá de los autores mismos.

La frecuencia absoluta que obtiene para cada vocablo se vuelve relativa en relación con el total de ocurrencias de vocablos en el corpus, para dar por resultado sus porcentajes de frecuencia. A base de estos datos, mediante cálculos estadísticos, logra definir una medida de la cercanía entre los vocablos, respecto de un "espacio semántico" n-dimensional, constituido por todo el vocabulario de su corpus y las diferentes relaciones entre vocablos que se crean a partir de cada uno de ellos. Como no es posible explicar aquí sus procedimientos, nos concretaremos a considerar algunos de sus resultados.

Tomando como punto de partida el vocablo *buch* 'libro', la lista de vocablos relacionados con él, en orden de mayor a menor relación de dependencia —con algunas discrepancias debidas al cálculo, que no explicaremos aquí— fue:

1	buch	'libro'	1.000
2	referieren	'reseñar'	.507
3	rezension	'reseña'	.493
4	lesen	'leer'	.301
5	redaktion	'redacción'	.206
6	kritik	'crítica'	.387
7	sammeln	'coleccionar'	.106
8	fund	'depósito'	.162
9	müssen	'deber'	.139
10	letzt	'último'	.113
11	bleiben	'quedar, restar'	.093
12	journal	'diario'	.206
13	zeitung	'periódico'	.181
14	zahl	'cifra'	.053
15	rundfunk	'radiodifusión'	.053
16	erscheinen	'aparecer, salir a la luz'	.089
17	erkenntnis	'conocimiento'	.073

Como se puede ver, los resultados muestran relaciones asociativas en principio, que sin embargo corresponden a la manera en que el sentido de un texto selecciona vocabulario y lo ordena. Hay que notar también que este procedimiento no se circunscribe a sustantivos, sino que considera por igual verbos y adjetivos (probablemente también adverbios). Llevadas estas medidas a representaciones gráficas, lo que se desprende de ellas es que el vocabulario de un texto o de cierto tipo de textos forma *conglomerados*, analizables topológicamente (*conglomerado* es la correspondencia española del inglés *cluster*, un vocablo propio tanto de la astronomía como de los análisis estadísticos). Es decir, que a partir de cada vocablo que tomemos como principio de un análisis en conglomerados, veremos que el resto de los vocablos en el texto se reagrupa, formando "nubes" o estructuras de significado.

5. CONCLUSIONES

Lo que se puede concluir de este capítulo es, por un lado, que los estudios cuantitativos del léxico ofrecen una ayuda muy importante para su estudio, así como para diversas aplicaciones orientadas a la enseñanza de una lengua o a la elaboración de métodos de diagnóstico de ciertos daños al oído o al cerebro; por el otro, que ofrecen una idea valiosa acerca de la manera en que el léxico opera en la memoria de las personas y en la "memoria colectiva". Aun cuando hoy en día hay una copiosa bibliografía de estudios cuantitativos del léxico, sus requerimientos de formación matemática han impedido que la mayor parte de los lingüistas podamos aprovecharla y tomar en cuenta sus hallazgos para comprender mejor el vocabulario.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

frecuencia absoluta y relativa; dispersión; uso; ponderación de un corpus; disponibilidad léxica; análisis en conglomerados.

OBRAS CONSULTADAS

GARCÍA HOZ, Víctor, *Vocabulario usual, común y fundamental. Determinación y análisis de sus factores*, Madrid, CSIC, 1953.

GOUGENHEIM, Georges, *et al.*, *L'élaboration du français fondamental (1er. degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, París, Didier, 1964.

JUILLAND, Alphonse, y Emilio CHANG RODRÍGUEZ, *A Frequency Dictionary of Spanish Words*, La Haya, Mouton, 1964.

LARA, Luis Fernando, y Roberto HAM CHANDE, "Base estadística del *Diccionario del español de México*", en L.F. LARA, R. HAM e I. GARCÍA HIDALGO, *Investigaciones lingüísticas en lexicografía*, México, El Colegio de México, 1980.

LÓPEZ CHÁVEZ, Juan, *¿Qué te viene a la memoria? La disponibilidad léxica: teoría métodos y aplicaciones*, Zacateca, Universidad Autónoma de Zacatecas, 2003.

MULLER, Charles, *Estadística lingüística*, Madrid, Gredos, 1973.

RIEGER, Burghard B., "Fuzzy Word Meaning Analysis and Representation in Linguistic Semantics. An Empirical Approach to the Reconstruction of Lexical Meanings in East- and West-German Newspaper Texts", en *COLING 80. Proceedings of the 8th International Conference on Computational Linguistics*, Tokio, ICCL, 1980, pp. 76-84.

——, "Generating Dependency Structures of Fuzzy Word Meanings in Semantic Space", en S. Hattori y K. Inone (eds.), *Proceedings of the XIIIth International Congress of Linguists*, Tokio, CIPL, 1983, pp. 543-548.

——, "Connotative Dependency Structures in Semantic Space", en B.B. Rieger (ed.), *Empirical Semantics*, Bochum, Brockmeyer, 1981, II, pp. 622-711.

9
EL ESTUDIO CUALITATIVO DEL SIGNIFICADO LÉXICO

INTRODUCCIÓN

Ahora reuniremos los conocimientos adquiridos en los capítulos 1, 2 y 4, especialmente en la parte que se refiere a los métodos de análisis estructural de los signos lingüísticos, con los métodos de abordaje explicados en los dos capítulos anteriores, para proceder al análisis semántico, *cualitativo*, del léxico.

Dijimos que los resultados del método de investigación del léxico disponible nos llevan a las siguientes conclusiones: *a*] que los resultados cuantitativos obtenidos nos demuestran que, aunque la memoria del léxico es un fenómeno individual, las memorias particulares de los miembros de un grupo social coinciden, en las medidas explicadas, y conforman lo que podemos llamar una "memoria social" del léxico, y *b*] que el procedimiento de descubrimiento de ese léxico disponible se basa en el más elemental fenómeno psicológico, que es el de la *asociación libre*.

En este capítulo tomaremos como base esos dos descubrimientos para desarrollar el estudio cualitativo del léxico, que nos llevará a una de las cuestiones más debatidas de la semántica léxica, que consiste en la probable existencia de una estructura del léxico de características similares a las estructuras fonológicas y morfológicas de la lengua.

1. CONCEPTOS BÁSICOS

En 1, § 1.3 explicamos el modo en que se utiliza el *método de la conmutación* entre signos lingüísticos para descubrir sus características. Decíamos que este método, que es ante todo un *método de descubrimiento*, se basa en la posibilidad de comparar dos signos entre sí. Para poder efectuar esa comparación es necesario que en ambos haya alguna caracterís-

tica que los haga comparables. Por ejemplo: si pretendemos comparar las palabras *edificio* y *hierba*, no podremos reconocer en ellas los fonemas que las componen (pues no permiten formar pares mínimos), e identificar sus posibles morfemas requeriría encontrar entre ellas alguna semejanza, que permitiera alguna segmentación; si tomáramos otras dos, de cuyos morfemas finales pudiéramos afirmar que tienen en común con ellas dos morfemas de género *-o*/*-a*, podríamos suponer que están compuestas por los morfemas *edifici-*, *hierb-*, *-o* y *-a*, aunque más tarde tendríamos que descubrir que la segmentación más adecuada de la primera palabra sería *edific-i-o*, al introducir en la comparación *edific-ar* y otras formas de la conjugación del verbo, así como tendríamos que proponer dos formas supletivas *hierb-* y *herb-* (pues no se dice *hierbario, sino *herbario*, no *hierbívoro sino *herbívoro*) (acerca del asterisco que precede estas formas, véase el capítulo 11). Para analizar su sustancia del contenido, como nos habría sucedido con su fonología, no encontraríamos ninguna semejanza de significado que nos permitiera oponerlas y, en consecuencia, no podríamos proceder al estudio de sus significados.

Para poder utilizar el método de la conmutación, en consecuencia, es necesario que, entre dos palabras que queramos analizar, haya algún elemento de significado común a ambas, que llamamos *base de la comparación* —el gran lingüista rumano Eugenio Coseriu, cuyo pensamiento influirá en este capítulo, la llamaba *valor de campo*—; todos los demás elementos de significado que encontremos diferentes entre ellas, serán *rasgos significativos* o *semas* (concebidos de manera paralela a la fonología, que descubre *rasgos distintivos*).

1.1. El campo asociativo

Y bien, entonces nos planteamos la pregunta de cómo reunir vocablos que nos permitan encontrar con facilidad suficientes bases de la comparación como para que el método de la conmutación sea aplicable.

En principio, se puede proceder a reunir un gran corpus de textos en los que podamos reconocer un tema central en común. Por ejemplo, textos literarios que traten las pasiones humanas, la inteligencia humana, los amores trovadorescos; textos médicos sobre los movimientos del cuerpo humano; textos populares acerca de la vida de los pájaros, o lo

que nos parezca más necesario para alguna investigación particular o más interesante. De esos textos se pueden entresacar los vocablos que requerimos, en general, sustantivos, adjetivos y verbos, según nuestro propio conocimiento de la lengua.

Otra manera de reunir vocablos para el estudio de su significado es el método del vocabulario disponible, pues si la encuesta se ha realizado proponiendo núcleos temáticos, por ejemplo, "animales salvajes", "muebles domésticos" o, de nuevo, "pasiones humanas", etc., es la memoria colectiva de nuestros informantes la que nos ofrece vocablos que puedan tener buenas bases de la comparación entre todos.

Por último, probablemente el método más objetivo para reunir vocablos y poder someter sus significados a un análisis semántico, sea el de *conglomerados*, como lo expusimos en el capítulo anterior a propósito de las investigaciones de Rieger, pues nos ofrece la ventaja de que no depende de nuestros conocimientos previos.

Pero lo común, que seguiremos haciendo por mucho tiempo, en tanto no se difundan los métodos cuantitativos, consiste en ampliar el procedimiento de asociación que da lugar al vocabulario disponible, invitando a nuestros informantes a que hagan un ejercicio de *asociación libre*, tal como acostumbra hacerse en psicología y en psicoanálisis. Un vocabulario recogido por asociación libre nos da un *campo asociativo*.

Un campo asociativo está formado por: *a*] vocablos relacionados por el tema o los objetos de la experiencia a los que refieren (incluso términos técnicos de una especialidad, conocidos por los informantes); *b*] vocablos con la misma etimología; *c*] vocablos agrupados por su categoría gramatical; *d*] vocablos agrupados por sus características morfológicas y fonológicas, y además, probablemente, *e*] vocablos de diferentes dialectos o variedades regionales o sociales de la misma lengua.

Tomemos el siguiente ejemplo: si a un hispanohablante le proponemos tomar la palabra *casa* como punto de partida de un proceso de asociación libre obtendremos una constelación de vocablos como los siguientes: cuarto, baño, sala, cocina, vivienda, choza, cabaña, palacio, castillo, chalet, departamento, edificio, techo, pared, ventana, puerta, jardín, huer-

ta, alberca, piscina, pasto, árbol, hogar, familia, padres, hermanos, abuelos, parientes, refugio, descanso, bienestar, calor, cobijo, resguardo, protección, cama, sofá, silla, sillón, cobija, sábana, almohada, cubiertos, cuchara, tenedor, cuchillo, hotel, ayuntamiento, oficina, casita, casucha, caserío, caserón, casero, masa, pasa, tasa, gasa, capa, cara, cala, etcétera.

Si seguimos el mismo procedimiento con varios informantes, encontraremos muchas coincidencias entre ellos, que nos pueden proporcionar el punto de partida de nuestro estudio léxico.

En esa constelación podemos organizar los vocablos mediante diferentes criterios: tipos de edificio, habitaciones de la casa, elementos de construcción, muebles domésticos, estructura familiar, sentimientos que inspira la vida doméstica, varios derivados con diferentes significados, y vocablos con expresión semejante, que nutren un estudio de pares mínimos para la fonología. Una vez que hemos hecho esta organización, podemos tomar los vocablos correspondientes a cada criterio y comenzar a analizar el léxico reunido.

1.2. Descubrimiento e interpretación

Durante varias decenas de años se ha discutido la objetividad del análisis semántico, tomando como punto de referencia la que nos ofrece el análisis fonológico. Pues siendo el fonema un fenómeno perceptual determinado por la naturaleza del sonido, los rasgos distintivos que descubre en él un lingüista, ya sea a base exclusivamente de conmutaciones de pares mínimos o con ayuda de estudios espectrográficos acústicos, adquieren una objetividad tal que se pueden comprobar sobre la base del sonido o de diversas pruebas de percepción y que, en principio, puede verificar cualquier otro lingüista. A pesar de ello, las diferencias que aparecen al final de los estudios fonológicos de una lengua no sólo son efecto de las necesidades de coherencia y de economía de la teoría final, sino del hecho de que los fonemas y sus pronunciaciones dependen de la pertinencia comunicativa que tengan para los miembros de la comunidad lingüística estudiada de la que, generalmente, el lingüista no forma parte.

La sustancia del contenido de los signos no tiene, a diferencia de la sustancia de la expresión, ningún anclaje natural. Es un fenómeno de la mente humana, respecto del cual incluso sus correlatos neuronales no

son de mucha ayuda para garantizar una objetividad del análisis. En tales circunstancias, la objetividad que se pueda alcanzar no es de la misma clase que la que produce la fonología, pues tanto los signos mismos pueden utilizarse de maneras ilimitadamente diferentes, reduciendo su significado, ampliándolo o modificándolo metafóricamente, como los significados obtenidos varían de acuerdo con sus contextos de uso, los conocimientos y la sensibilidad del lingüista que realiza el análisis.

Ello nos lleva al convencimiento de que el análisis semántico del léxico es un *método interpretativo*; lo que se descubre al tomar en consideración los vocablos depende tanto del contexto en que se utilizan como de la experiencia, el conocimiento y la sensibilidad lingüística del analista. Pero que sea un método interpretativo no lo vuelve necesariamente un capricho subjetivo, pues el método tiene como una de sus metas principales garantizar una objetividad tal que sus resultados sean *intersubjetivamente verificables* y den lugar a una verosimilitud ampliamente aceptable.

1.3. La preparación del análisis

En consecuencia, para poder proceder al análisis léxico es necesario delimitar con cuidado la clase de datos que se habrán de analizar, con el objeto de aclarar y despejar algunos factores que intervienen en la realidad de la significación y que pueden interferir gravemente en los resultados; a esta delimitación previa la llamaremos *preparación del análisis*.

La preparación del análisis consta de: *a*] una delimitación del estado de lengua del que forma parte el léxico en estudio; *b*] claridad acerca de las diferentes realidades dialectales que pueden confluir en el léxico bajo estudio; *c*] claridad acerca de la muy probable y real mezcla de tradiciones verbales en el léxico, y *d*] una distinción clara entre *significado* y *designación* en el modo de usar los vocablos.

1.3.1. Estado de lengua y sincronía

La metodología general de la lingüística estructural exigía que el léxico que se fuera a analizar correspondiera a un mismo estadio histórico de la lengua; es decir, que no se entremezclaran vocablos correspondientes

a diferentes épocas de la lengua histórica. Por ejemplo, si se tratara de analizar el léxico contemporáneo del calzado, no podrían tomarse como elementos del estudio vocablos como *borceguí* o *polaina*, pues probablemente ya no correspondan al español actual y poca gente conozca su significado; si quisiéramos estudiar el léxico de los barcos, no se podría considerar vocablos como *bergantín*, *carabela*, *galera*, *nao* o incluso *brulote* (que era una nave incendiada que empujaban en contra de otra durante un asalto naval), que son nombres de barcos existentes en la antigüedad, pero que ya no se usan hoy en día; y aun en el caso de *fragata*, una palabra que se sigue utilizando aunque su significado ha cambiado junto con las características de las fragatas actuales, es necesario tomar en consideración contextos de uso de la palabra correspondientes únicamente a la época actual.

Un *estado de lengua*, es decir, un momento histórico de una lengua, del cual se pueda suponer una contemporaneidad de todos los datos que permita postular la sincronía de su sistema, se puede definir mediante diferentes criterios; uno de ellos, el de más amplia aplicabilidad, consiste en tomar en cuenta qué generaciones de seres humanos conviven en la misma época y se comunican entre sí. Por ejemplo, hoy en día conviven generaciones nacidas a finales de 1920 con generaciones de niños nacidos hace siete años (en general se considera que un niño de siete años de edad domina ya el léxico fundamental de su lengua materna) y todas las intermedias. Abuelos, incluso bisabuelos, padres e hijos hablan entre sí y dominan buena parte de los significados de los signos contemporáneos de su lengua materna. Se puede por eso considerar que el *estado de lengua* "contemporáneo" cubre de 1920 a 2010.

Es muy importante delimitar con claridad el estado de lengua que se vaya a estudiar, para poder seleccionar adecuadamente los datos; sin embargo, cuando en ese estado se llegan a utilizar vocablos correspondientes a otros estados anteriores, aunque sus significados hayan cambiado (por ejemplo, hoy en día casi nadie distinguirá entre una carabela y una nao), *tienen un significado* para sus hablantes y hay que analizarlo (de *carabela* y *nao* muchos hablantes de español dirán que significan "barcos de vela, de la época del descubrimiento de América"; ya no se encontrarán diferencias entre ambos significados, pero eso no quiere decir que no tengan algún significado).

Los vocablos reunidos correspondientes al estado de lengua pasan a considerarse elementos de la misma *sincronía*, es decir, elementos que forman parte de las estructuras vigentes en el sistema de la lengua.

1.3.2. La unidad dialectal

Todos los seres humanos nacemos y vivimos en algún lugar, que se separa de los demás por la distancia, por obstáculos como las montañas, los ríos o los mares, por la clase de actividades laborales de sus habitantes, por las relaciones que establecen con localidades vecinas, etc. En cada localidad se produce naturalmente una *comunidad de comunicación*, es decir, un conjunto de valores, intereses, acontecimientos, actividades, etc. que comparten más o menos todos sus habitantes. Es probable que cada comunidad de comunicación se caracterice también por sus maneras de hablar y, en el caso del léxico, por el vocabulario que utilice. El análisis léxico debe tomar en cuenta los datos correspondientes a esa comunidad. Cuando lo que se enfoca es la *misma lengua* —lo que, en realidad, es una abstracción sistemática, simbólica e ideológica— todas las comunidades de comunicación en que se habla se convierten en *dialectos* de la lengua. Es decir, se habla de dialectos cuando se comparan formas de hablar de localidades o grupos sociales diferentes de la misma lengua. En consecuencia, los datos que se han de analizar conviene que correspondan a un mismo dialecto; si no —que es lo que ocurre hoy en día, como efecto de los medios de comunicación— hay que tratar de reconocer las mezclas dialectales para tomarlas en cuenta, en caso necesario, durante el análisis.

1.3.3. Las tradiciones verbales

En toda sociedad la lengua se utiliza para una enorme cantidad de contactos humanos y mensajes. Se habla para saludarse, para entablar una conversación, para compartir instrucciones, para dar órdenes, para participar en actividades de grupo, para practicar una religión, para aprender en la escuela, para jugar, para enamorar, para dar pésames y manifestar compasión por los familiares de los difuntos, para

dirimir juicios, para hacer política, etc. Cada comunidad desarrolla maneras de hablar propias, correspondientes a sus experiencias en el uso de la lengua, y esas maneras de hablar se enseñan de padres a hijos dando lugar a una *tradición verbal*. En muchas comunidades se llegan a separar diferentes maneras de hablar la lengua según necesidades que se consideran importantes; por ejemplo, entre los pueblos antiguos que conocemos, el modo de hablar de lo sagrado, es decir, el modo de hablar de los dioses o de Dios, las oraciones, los ensalmos, los hechizos y la conservación de las palabras sagradas daban lugar muchas veces a un vocabulario especial y a unas fórmulas morfológicas y sintácticas diferentes de las del modo de hablar ordinario; en muchas comunidades amerindias, diferentes grados de cortesía que exige la jerarquía social producen vocabularios diferentes. Esos usos cultivados a lo largo de los siglos, y todavía transmitidos de una generación a otra, originan las *tradiciones cultas*. En lenguas como el español, que tienen un sistema de escritura y una tradición milenaria de escritura, la *lengua escrita* o *literaria* —de allí viene la palabra *literatura*: de lo que se escribe con letras o *litterae* en latín— es el origen de su tradición culta.

Pero junto con esas tradiciones, las experiencias particulares de las comunidades locales, de los pueblos, de las regiones, incluso de las naciones, que se transmiten al calor del contacto directo y espontáneo y no llegan a documentarse literariamente, dan lugar a varias *tradiciones verbales populares*. Hablando de México, por ejemplo, esa manera de hablar que oímos en las películas de mediados del siglo XX, dedicadas a temas y personajes de barrios populares, entre los que surgieron cómicos como Cantinflas, Mantequilla, Borolas o Harapos, era reflejo de una tradición verbal popular de los barrios del centro de la Ciudad de México, entre los que destacó Tepito. Pero lo mismo se puede decir del hablar madrileño del barrio de Lavapiés o del bonaerense de La Boca.

Todos los hablantes de una lengua solemos utilizar ambas tradiciones e incluso mezclarlas. Para preparar el análisis léxico es necesario tener en cuenta que vocablos originados en ambas tradiciones pueden aparecer mezclados, si queremos interpretar adecuadamente sus diferencias de significado.

1.3.4. Significación y designación

Una lengua es un instrumento de significación; es decir, un instrumento mediante el cual manifestamos a las personas que nos rodean nuestras experiencias de la vida, sean cuales sean, concretas o abstractas, buenas o malas, más sensuales o más intelectuales, verdaderas o falsas, reales o imaginadas. Lo que hace de la lengua el mejor instrumento de significación de que disponemos los seres humanos es un sistema lingüístico compartido por todos los que la hablan y un acervo común de memorias de esas experiencias cristalizado en los significados de los signos lingüísticos.

Podemos significar nuestras experiencias de la vida de varias maneras: podemos nombrarlas, utilizando el significado común de las palabras: el significado que todos los miembros de la comunidad reconocen de inmediato. Por ejemplo: si enunciamos aisladamente la palabra *gorila* entenderemos que significa "mono de gran tamaño, fuerte y en muchos casos feroz", y después comprenderemos gracias a ese significado que alguien, cuando vea arremeter contra una persona o un grupo de personas a unos hombres fuertemente armados, amenazadores y de gran corpulencia, grite: "¡cuidado, que vienen unos *gorilas*!"; en cambio, al oír en aislamiento esa palabra, fuera de contexto, no entenderíamos inmediatamente 'hombre corpulento, de aspecto feroz y amenazador'. Cuando usamos la palabra *gorila* con ese segundo significado reconocemos que se trata de un significado producto de una metáfora. Podemos significar algo a base de un predominio de metáforas, que lleguen a producir en nosotros un efecto sorpresivo, muchas veces deslumbrante y estético, como sucede en la poesía. Cuando Luis de Góngora escribía, en su poema de "La Galatea": "Era Acis un venablo de Cupido", si sabemos que Acis era el nombre de un joven, de quien estaba enamorada Galatea, entenderemos que *venablo* está usado metafóricamente, pues un ser humano no puede ser una flecha disparada por el diosecillo del amor.

Pero también podemos tratar de significar algo con extrema precisión y hacer que el nombre que le pongamos sólo signifique ese algo y no otra cosa. En ese caso, significamos *designando* nuestro objeto. En la comunicación técnica y científica, sobre todo, nos esforzamos porque la *designación* de los objetos científicos, de sus procedimientos, de sus métodos y sus teorías nos asegure un discurso objetivo, que permita una comunicación unívoca y eficaz que dé lugar a una verificación. Podemos

decir que el uso de una lengua para designar algo se orienta al eficaz y preciso reconocimiento de ese algo; en tanto que a nuestros usos ordinarios de la lengua les basta lograr una comprensión suficiente, pero no necesariamente tan objetiva y precisa de lo que queremos decir como sucede en la comunicación especializada.

La designación es uno de los modos de la significación.

La designación da lugar a las nomenclaturas; en particular, a los términos científicos y técnicos. Tomemos por caso la palabra *ceiba:* muchos hispanohablantes, sobre todo los que vivimos en los trópicos, entendemos que se trata de un gran árbol, cuya copa se eleva mucho sobre su tronco y adquiere la forma de agrupaciones horizontales de ramas, y tiene su tronco tachonado de protuberancias relativamente espinosas; pero hay varios árboles llamados *ceiba* en las costas del Golfo de México, en Yucatán, en Centroamérica, etc. que tienen diferencias entre sí, notables para el botánico. Para poder reconocer con precisión el *objeto* (o los objetos) llamado *ceiba*, tenemos que acudir a los botánicos, quienes nos ofrecerán una nomenclatura linneana de las ceibas, como *Ceiba pentandra*, y una descripción como "árbol muy grande de la familia de las bombacáceas, de hasta 40 metros de altura, de tronco con contrafuertes desde la base y corteza tachonada con grandes formaciones cónicas, consideradas espinas; copa muy amplia y extendida; sus hojas son palmeadas, compuestas de 5 a 7 hojillas oblongas; da flores blancas o de color rosado; su fruto es largo y elíptico y contiene semillas grandes, envueltas en fibras algodonosas". Esa designación científica de una de las especies de árboles llamadas *ceiba* (*Ceiba pentandra*) y su descripción botánica forman parte del conjunto de los términos técnicos botánicos. Los rasgos significativos que distinguen a los árboles de esa familia para la botánica no son rasgos lingüísticos, es decir, no son rasgos formados en la comunicación ordinaria de las comunidades que tienen contacto con las ceibas, sino que son rasgos determinados por la taxonomía botánica. En cambio, entre varios pueblos herederos de la cultura maya, la ceiba no sólo se reconoce físicamente, sino que se considera el puente entre la tierra y el cielo, de gran valor cosmogónico, por lo que es posible que en los significados de los vocablos que las nombran en maya yucateco, en

tzeltal, en tzotzil, en tojolabal, en chol, en maya quiché, etc. haya rasgos significativos determinados por esa comprensión social del árbol.

Tomemos un ejemplo más: para cualquier hispanohablante, *agua* quiere decir "líquido transparente que se bebe para apagar la sed; se utiliza para lavarse y para regar y fertilizar la tierra; se produce en las nubes, de las que cae en forma de lluvia o se encuentra en mares, lagunas, ríos, etc.; por lo general no tiene sabor, aunque la de los mares es salada". Para un químico, los componentes necesarios del agua son las moléculas formadas por dos átomos de hidrógeno y uno de oxígeno, aunque también podrá incluir los diferentes minerales que suele arrastrar con ella. Para el químico el vocablo *agua designa* a la molécula de H_2O; para el químico, el vocablo *agua* es un término técnico con que se *designa* un objeto de la naturaleza.

Es necesario, en consecuencia, tomar en cuenta que un análisis semántico de un vocabulario se interesa ante todo por el significado de los vocablos, no por los objetos que designan (pues, como dijimos antes, la designación es solamente una manera de significar). Es decir, que el análisis semántico de los vocablos de una lengua, cuando se orienta al conocimiento de esa lengua, es un análisis del significado social u ordinario de los signos; en tanto que cuando se orienta al estudio de un vocabulario científico o técnico, es un análisis de la designación de objetos considerados por la ciencia o la técnica; y hay que cuidar que no se confundan rasgos significativos para la ciencia o para la técnica de los objetos que estudian, con rasgos significativos de la lengua.

En la química, los nombres de los elementos naturales son designaciones convencionales de las diferentes materias descubiertas en el universo. Ninguno de esos nombres ha adquirido un significado diferente en la lengua histórica, por lo que, cuando una persona habla de *sodio*, *potasio*, *gas xenón* o *vanadio*, está utilizando esos signos como términos de la química y, en consecuencia, su significado no se puede analizar como si correspondiera a la lengua ordinaria. Esos nombres son signos lingüísticos, pero sólo significan designando elementos. Lo que le da "estructura" a esa nomenclatura química no es un significado socialmente acuñado, sino las propiedades atómicas de los elementos. Ello no impide que algunos de ellos adquieran significado para una sociedad. Cuando alguien se desmaya y lo rodea la gente, alguien puede pedir: "¡déjenlo respirar, necesita *oxígeno*!", pues se entiende que el oxígeno es 'aire'.

Los términos especializados son como los nombres propios: por ejemplo, el signo *Juan* no tiene un significado, sino que sólo identifica a las personas a las que se bautiza o se registra de esa manera. (En relación con el significado etimológico de los nombres propios, véase el capítulo 11.) Sólo en el siglo XIX mexicano y durante la Revolución de 1910, cuando los ejércitos de los diferentes caudillos campeaban por el país, la gente optó por llamar "juan" a cualquier soldado, por lo que el signo *juan* pasó a querer decir 'soldado', en particular 'soldado raso'.

Muy probablemente se den cuenta de que lo tratado en los cuatro parágrafos anteriores contradice la enseñanza estándar de la semántica estructural, en especial la de Eugenio Coseriu; es consecuencia de una superación crítica del estructuralismo y un cambio de visión hacia la lingüística pragmática.

2. DEL CAMPO ASOCIATIVO AL CAMPO LÉXICO

Una vez terminada la preparación del análisis, podemos proceder a tomar en cuenta vocablos de un mismo tema, de una misma categoría gramatical, para buscar una base de la comparación. Tomemos la serie de vocablos agrupados como tipos de edificio: *casa*, *edificio*, *departamento*, *choza*, *cabaña*, *chalet*, *palacio*, *castillo*. No son los únicos vocablos existentes en español, pero sí los que aparecieron en nuestro experimento del campo asociativo. Entre todos ellos podemos proponer como base de la comparación o valor de campo el rasgo 'habitación construida por seres humanos'. Si hubiéramos propuesto como base de la comparación el rasgo 'habitación de seres humanos' lo habríamos hecho para poder incluir *cueva* y *caverna*.

2.1. Las características de un sema

Una de las adquisiciones irrenunciables del estructuralismo es el *análisis binario*. Éste consiste en reducir todas las diferencias que encontremos entre diferentes fonemas o entre diferentes signos a una oposición

binaria de sí o no. Cuando la oposición se reduce a esa distinción, podemos estar seguros de que no ha quedado nada fuera de nuestro análisis; que hemos llegado a una oposición *mínima*.

Un rasgo distintivo en fonología o significativo en semántica debe ser el resultado de un análisis binario. Por ejemplo, entre los fonemas /p/ y /b/ del español, la oposición entre ambos se reduce a la existencia o no del rasgo "sonoridad".

El análisis binario, sin embargo y a diferencia de lo que ocurre con el análisis del sonido o de la imagen, que se puede ir reduciendo infinitesimalmente, como se ve en las cámaras fotográficas, cuya sensibilidad medida en pixeles (*picture cells* 'celdillas de imagen', otro acrónimo) se amplía año con año, en el análisis lingüístico tiene un límite: la *pertinencia* del rasgo en relación con el campo que se está analizando.

El *binarismo* fue el punto de partida de la invención de la computadora electrónica que, en último análisis, es una enorme serie de *switches* o apagadores, que dejan o no pasar la corriente eléctrica por sus circuitos. Los "dígitos binarios" o *bits* —acrónimo de *binary digit*, una feliz creación terminológica que aprovecha el significado del vocablo inglés *bit* 'pedazo'— son representaciones mediante números o dígitos (cero y uno) de esa clase de análisis binarios, hoy extendidos a la fotografía, la grabación de discos musicales, la televisión, el teléfono, etcétera.

Supongamos que quisiéramos analizar los significados de los vocablos *cueva*, *caverna*, *antro*, *cubil* y *gruta*, lugares todos en donde los seres humanos pueden guarecerse e incluso adaptar como habitaciones; en relación con ese campo, el rasgo 'habitación construida' no sería pertinente, en cambio, lo sería 'refugio natural', por ejemplo.

Un rasgo significativo debe ser, por lo tanto, mínimo y pertinente. Tal es la definición del término técnico *sema*.

2.2. El procedimiento de análisis

Con esa preparación, podemos comenzar nuestro análisis del significado de los vocablos que hemos agrupado bajo la misma base de la comparación. En una tabla pondremos, en la columna izquierda, el vocabulario en estudio; en las columnas sucesivas, marcaremos con los símbolos + y – la existencia o presencia de los semas que nos parezcan pertinentes. En la fila superior, el enunciado del significado de esos semas. El sema 1 será siempre el de la base de la comparación.

sema (s)	habitación construida (s 1)	para una familia (s 2)	con entrada particular (s 3)	agrupada con otras en la misma construcción (s 4)	de estilo alpino (s 5)	de material provisional (s 6)	de troncos o madera (s 7)	de la aristocracia o el gobierno (s 8)	con sistema defensivo (s 9)
casa	+	+	+	-	-	-	-	-	-
departamento	+	+	-	+	-	-	-	-	-
edificio	+	-	+	+	-	-	-	-	-
choza	+	-	+	-	-	+	+	-	-
cabaña	+	-	+	-	-	-	+	-	-
chalet	+	+	+	-	+	-	+	-	-
palacio	+	-	-	-	-	-	-	+	-
castillo	+	-	-	-	-	-	-	+	+

En la interpretación que hemos seguido, a partir del significado de los vocablos en el español mexicano, hemos propuesto nueve semas para establecer las distinciones de significado de cada vocablo y los hemos numerado.

2.2.1. El semema

Llamamos *semema* al conjunto de semas en que hemos descompuesto el significado de cada vocablo. Lo representaremos mediante la fórmula S_x (s 1, s 2, ..., s n). Consideraremos semas de cada semema solamente los positivos. En nuestro ejemplo, los sememas de los vocablos correspondientes son:

casa S (s 1, s 2, s 3)
departamento S (s 1, s 2, s 4)
edificio S (s 1, s 3, s 4)
choza S (s 1, s 3, s 6, s 7)
cabaña S (s 1, s 3, s 7)
chalet S (s 1, s 2, s 3, s 5, s 7)
palacio S (s 1, s 8)
castillo S (s 1, s 8, s 9)

En estos resultados podemos comprobar: *a*] que la base de la comparación se encuentra en todos los sememas, por lo que es un valor de campo; *b*] que los semas propuestos sirven para establecer distinciones entre los significados de los vocablos estudiados; *c*] que bastan los nueve semas propuestos para establecer distinciones binarias entre los significados de los vocablos. (Otros semas, que podrían proponerse, si resultaran más pertinentes que los considerados, tendrían que sustituir a éstos; si no, resultarían redundantes y serían por ello innecesarios.) Ello no elimina la posibilidad de adoptar otro criterio de interpretación, que conduzca a otros semas, pero siempre tendría que verificarse su adecuación al conjunto de vocablos considerados, sometiendo a nuestros informantes los sememas obtenidos y preguntándoles si les parecen aceptables y verosímiles para representar el significado de cada vocablo.

2.2.2. La formulación del sema

Los dos problemas centrales del análisis estructural que no llegaron a solucionarse debido al enfoque rígidamente positivista de esa corriente del pensamiento lingüístico, fueron, por un lado, el de la manera de representar o formular los rasgos; por el otro, el del probable carácter sistemático de los semas y los sememas obtenidos.

Pues la semántica estructural esperaba que los semas obtenidos correspondieran a una realidad significativa de todas las lenguas y que, en consecuencia, una vez "descubiertos", fueran válidos para analizar con ellos cualquier conjunto de vocablos de cualquier lengua (tal era la idea que se echa de ver en los estudios pioneros del lingüista francés Bernard Pottier, y que se descubre también en las investigaciones seudoformalistas de los lingüistas Igor Melchuk y Anna Wierzbicka). Los semas serían entonces entidades independientes de cada lengua y, en vez de ser resultado de una *descomposición interpretativa del significado*, serían elementos de una gran estructura sémica universal, que cada lengua *realizaría* —una noción querida por el estructuralismo— de diferente manera; es decir, que cada lengua seleccionaría semas de acuerdo con sus necesidades significativas. Al tratarse de una cantidad universal de semas, debería ser posible manejarlos mediante una especie de álgebra lingüística, independiente de cualquier lengua. De haber sido así, los semas tendrían para la semántica un papel semejante al de los átomos en la física y, en consecuencia, habría una estructura sémica universal, de la que se desprendería la estructura sémica del léxico de cada lengua.

La realidad es que el significado de cada sema, construido como un lenguaje de descripción, adquiere su sentido en la misma lengua en que se realiza el análisis y se puede formular de muchas maneras diferentes, de acuerdo con la reducción que hace el semantista para lograr unidades mínimas pertinentes. Los semas no son entidades independientes de la lengua en que se hace el análisis, ni existen como partes de un elenco gigantesco, que el análisis se limite a descubrir y que valga para cualquier lengua.

La descomposición del significado en semas es una técnica muy útil y necesaria para explorar la conformación del significado de los vocablos, pero no es un procedimiento de descubrimiento, sino de interpretación. La "estructura" sémica de los vocablos agrupados en un conjunto léxico como el estudiado es efecto de los criterios interpretativos del análisis y, para cada conjunto, puede corresponder a la estructuración parcial que tenga vigencia en una comunidad lingüística dada, en una época determinada y bajo las condiciones que impone el contexto de la comunicación. En la medida en que la estructuración construida corresponda a la memoria social del léxico, como sucede con las definiciones de los diccionarios, en esa medida podemos considerar que existen, efectivamente, estructuraciones del significado léxico.

Hoy en día conviene seguir la tesis de que no hay una gran estructura léxica de cada lengua y mucho menos una estructura sémica universal, sino que, por el contrario, las necesidades significativas de cada discurso o de cada texto organizan el vocabulario de acuerdo con sus necesidades, dotándolo de una estructuración. En vez de una estructura léxica o de un sistema léxico, hay que pensar que el léxico de una lengua forma una red, en la que se crean microestructuras según las condiciones del estado de lengua considerado, de la memoria social del léxico de ese estado y del contexto puntual en que se produce la significación.

2.3. El campo léxico

Llamaremos *campo léxico* al conjunto de vocablos que, sobre una base de comparación, podemos someter a juegos de oposiciones binarias que nos permiten asignar a cada uno de ellos un semema, relacionado con los otros sememas mediante semas compartidos. Muchos lingüistas llaman al campo léxico "campo semántico"; aquí no seguiremos esa costumbre porque la semántica no se reduce al léxico.

2.3.1. Las reestructuraciones de los campos

Una comprobación de la inexistencia de una gran estructura semántica del léxico es el hecho de que, si ampliamos los vocablos reunidos para un análisis de campo, los semas que tendremos que proponer serán relativamente diferentes. Ya veíamos antes lo que sucedería si al vocabu-

lario en estudio agregáramos *cueva*, *caverna*, *cubil*, *antro* y *gruta*: la base de comparación tendría que cambiar para poderlos incluir: una base de 'habitación' nos bastaría también para ello, pero tendríamos que agregar semas que sirvieran para distinguir *cueva* de *caverna*, de *cubil*, de *antro* y de *gruta*, que probablemente alterarían los sememas del vocabulario antes analizado. Tendríamos que agregar un sema 'para humanos', que distinguiera a todos los vocablos de *cubil*, que significa a una habitación de lobos y otras fieras salvajes; ese sema 'para humanos' o '+ humano' sería de tal generalidad, que nos llevaría a integrar el campo léxico de la habitación construida para personas a todo el conjunto de vocablos cuyo significado implique 'para humanos'. A la vez, tendríamos que agregar semas para distinguir la profundidad de cuevas, cavernas y grutas; un sema de 'con más de un habitáculo', que nos serviría para distinguir *gruta* de *caverna* y *cueva*, se colaría a los sememas de *casa*, *departamento*, *chalet*, *palacio* y *castillo*, para distinguirlos de *choza* y *cabaña*. Nuestro campo inicial se reestructuraría.

Este fenómeno se produce cada vez que se reúne una cantidad de vocablos para someterlos a análisis y formar con ellos un campo léxico. El campo léxico es, en consecuencia, una *microestructura* formada por el valor de campo escogido, dependiente del contexto y del conocimiento y la sensibilidad del analista. No hay pruebas de la existencia real de una gran estructura del léxico; en cambio, la capacidad humana de significar sus experiencias de la vida e interpretarlas da lugar a la formación de microestructuras léxicas, que no quedan fijas en la memoria, sino que son dinámicas y capaces de formar nuevas microestructuras.

2.3.2. Clasemas y marcadores semánticos

La semántica estructural, incluso la que se desarrolló durante los primeros 20 años de la corriente lingüística llamada "gramática generativo-transformacional", propuso que aquellos semas que puede uno atribuir a una gran cantidad de vocablos de muy diferentes campos, como es el caso del sema 'para humanos', se consideraran *clasemas* o *marcadores semánticos*. Suponen la existencia de grandes estructuras del significado léxico, a la vez que un intento por clasificar el vocabulario de tal

modo que se prestara para su *inserción* en el componente generativo-transformacional de la gramática. No discutiremos aquí tal concepción de la lengua.

Junto con esas unidades de análisis, el estructuralismo dio lugar a varias entidades más, que no enseñaremos aquí, por corresponder a la concepción general del sistema y del léxico propia de esa corriente, pero que no tienen sustento fenomenológico, ni se justifican del todo como instrumentos de análisis.

3. OTRAS MICROESTRUCTURAS LÉXICAS

Los métodos de estudio cuantitativo del léxico y el análisis binario del campo léxico nos han llevado a la posibilidad de poner en relación unos vocablos con otros, en vez de tener que suponer que las palabras son signos lingüísticos dispersos en la memoria y atenidos exclusivamente a aquello a lo que refieren de la experiencia de la vida. Incluso podemos afirmar que el método de la conmutación nos ha permitido reconocer unas palabras con otras; es decir, reconocer una existencia del significado de los vocablos en términos de relaciones con otros vocablos. *Hay léxico* y no meras sucesiones de vocablos en la memoria y en los datos, porque entre las palabras se forman redes de significado sobre la base del acervo histórico de la lengua y en dependencia de las necesidades de significación de cada comunidad, de cada persona y de cada contexto. En lo que sigue nos ocuparemos de las microestructuras léxicas más evidentes.

3.1. Los sinónimos

Es común pensar que vocablos como *ver* y *mirar*, *tirar* y *jalar*, *dar* y *otorgar*, *frijol* y *judía*, *betabel* y *remolacha*, *pavo* y *guajolote*, etc. son *sinónimos*, es decir, "que significan lo mismo". La continua edición de muchos "diccionarios de sinónimos" parece ser prueba suficiente de la existencia de esa clase de relación entre los vocablos de una lengua. Aquí trataremos de precisar qué clase de relaciones léxicas son las que subyacen a la noción de sinonimia.

3.1.1. Sinonimia referencial

Como vimos en los capítulos 2 y 4, las acciones significativas de los seres humanos se producen siempre en condiciones irrepetibles, que hacen que cada persona comience a formarse su comprensión del significado de cada palabra de una manera individual, pero en condiciones de diálogo con quienes la rodean, con lo que ajusta su comprensión a lo que los demás reciben como inteligible. Podemos imaginar la comprensión de los significados de las palabras en una gama continua, que va de lo individual a lo completamente ajustado a la lengua histórica, manifiesto en sus diccionarios. A lo largo de esa gama, naturalmente, se producen aparentes igualdades entre los significados de varias palabras.

La relación de sinonimia entre dos o más palabras que se nos hace más evidente es la que toma como base de significación una *referencia* a la experiencia de la vida; es decir, no revela una *igualdad de los significados* de las palabras consideradas sinónimas, sino una *equivalencia de esos signos respecto de lo referido*. Por ejemplo, los mexicanos entendemos como "sinónimos" las palabras *gis* y *tiza*, porque *nombran el mismo objeto*, que es el fragmento de yeso que sirve como instrumento para escribir en una *pizarra* o un *pizarrón*. *Tiza* es un vocablo que entendemos como característico del español nacional de España (aunque su origen sea náhuatl), en tanto que *gis* es característico del español nacional de México (aunque su origen sea griego). Lo mismo podemos decir de *pizarra* y *pizarrón*. Son sinónimos como nombres de objetos. La sinonimia, considerada en relación con objetos de la experiencia de la vida, es un proceso muy común de la significación. En este caso corresponde al uso de los signos como meros nombres de las cosas.

Sucede lo mismo con *betabel* y *remolacha*, *frijol* y *judía* o *pavo* y *guajolote*: en relación con aquello que nombran, son sinónimos, pero no lo son en cuanto a la manera en que se usan, pues la palabra *betabel* es característica del español mexicano, en tanto que *remolacha* se considera del español culto internacional (aunque habría que comprobarlo); *frijol* y *judía* sólo son sinónimos cuando se ha reconocido la clase de objetos a la que refieren, aunque los significados de ambas palabras difieran, pues el frijol mexicano es más pequeño que la judía, y la clase de la que forman parte es diferente en ambos dialectos del español. En español de México se distinguen los frijoles negros de los bayos; se cocinan refritos

o charros; ningún mexicano se atrevería a hablar de "judías refritas", temiendo que tal expresión, que involucra totalmente el signo, no tenga sentido en el otro dialecto. *Pavo* es un vocablo del español culto internacional, en tanto que *guajolote* forma parte de los dialectos mexicanos del altiplano central, puesto que en otras regiones de México a ese animal se le nombra *pípil*, *cócono* o *chumpipe*. Por las colocaciones comunes del español mexicano, se distinguen los usos de *pavo* y *guajolote*: *se come pavo* en Navidad, pero se hace *mole de guajolote*.

El lingüista alemán Klaus Heger proponía que dos o más vocablos podían estar en relación de sinonimia si y sólo si coincidían en sus respectivas tres funciones bühlerianas del signo: *referencial* (ordenada a los objetos y las relaciones entre objetos), *sintomática* (ordenada al hablante) y *apelativa* (ordenada al oyente).

Karl Bühler, un famoso lingüista y psicólogo austriaco, expuso de esa manera las funciones fundamentales del signo lingüístico en su libro *Teoría del lenguaje*.

Como se puede colegir, en todos los ejemplos considerados hasta ahora hay *sinonimia referencial*, pero no *sintomática*, pues esos signos revelan siempre al oyente la procedencia dialectal de su interlocutor. Una sinonimia en la función apelativa es todavía más difícil que se produzca, salvo en acciones en que el oyente desconozca la función referencial y se deje subyugar por el efecto apelativo del signo, como es el caso de un famoso ejemplo del siglo XIX, relatado por el lingüista Hugo Steinthal, según el cual un estudiante universitario logró insultar a una verdulera recitándole el alfabeto griego, que ella ignoraba, y, por lo tanto, resultó "sinónimo" de los peores insultos conocidos por ella.

3.1.2. La hiperonimia

Consideremos ahora las parejas de verbos ejemplificados: aunque muchos contextos nos demuestran que *ver* y *mirar* se usan como sinónimos, el significado de *ver* tiene valor de campo para todos los verbos que significan acciones de la vista, como *mirar*, *observar*, *divisar*, *vislumbrar*, *columbrar*, *espiar*, etc. *Mirar*, a su vez, podría tener valor de campo para *observar* y *espiar*, pero no para los demás. *Ver* significa 'percibir con la vista los objetos', por lo que puede ser el primer sema en un estudio de ese campo léxico, en tanto que *mirar*, que significa 'ver alguna

cosa con atención, dándose cuenta de ella', tiene un significado más restringido. Pero precisamente porque *ver* puede usarse como sustituto de todos los verbos pertenecientes a su campo léxico en determinados contextos, se llega a considerar, equivocadamente, sinónimo de todos ellos. La relación que se establece entre el significado de un vocablo con valor de campo para otro y que, por eso, sirve como sustituto del otro en determinados contextos, se llama *hiperonimia*.

Es un *hiperónimo* todo vocablo cuyo significado, por ser más general, puede sustituir parte del significado de otros vocablos, ayudando a su comprensión, aunque perdiendo precisión. Verbos como *ver*, *hacer*, *dar* pueden servir como hiperónimos de muchos otros, sin que sean sinónimos suyos; lo mismo se puede decir de vocablos como *cosa* y *objeto*.

Por último, consideremos la supuesta sinonimia entre *tirar* y *jalar*. En efecto, cuando una persona se encuentra delante de una puerta en un lugar público español, un letrero fijado en ella le ordenará *tirar*; en un lugar público mexicano, el vocablo correspondiente será *jalar*. La acción de atraer la puerta hacia uno mismo para poderla abrir es la misma y, de nuevo, referencialmente en ese contexto, *tirar* y *jalar* serán sinónimos. Sin embargo, mientras que en el español nacional de España, el significado de *tirar* como 'atraer algo una persona' es el predominante en ese contexto (aunque la Real Academia lo liste como la 24ª. acepción de su entrada en el diccionario), en el de México ese significado es el principal de *jalar*, pues *tirar* significa la acción opuesta, de 'hacer moverse alguna cosa en dirección opuesta a quien lo hace', aunque se entienda que en construcciones como *tirar de* (tirar de una carreta), el significado de *tirar de* es de atracción.

Podemos concluir entonces que la relación de sinonimia no es fija, no forma una estructura permanente de la lengua, sino que también se constituye en el contexto. Los "diccionarios de sinónimos", si no están bien hechos, confunden a sus lectores disipando la precisión del significado y haciéndoles creer, con ello, que las lenguas tienen una gran cantidad de sinónimos. Los buenos diccionarios de esa clase, como el de Samuel Gili y Gaya publicado por la editorial Vox, consisten más bien en una cuidadosa exposición de las diferencias que hay entre palabras de significados afines.

3.2. Los antónimos

Una relación de antonimia consiste en la existencia de dos o más vocablos cuyos significados se oponen simétricamente, es decir, que el semema del segundo significa la negación del primero: S/~S (a diferencia de lo que ocurre entre vocablos de un mismo campo léxico, en que no es negativo el semema, sino alguno de sus semas). Serían antónimos, por ejemplo, *sí* y *no*, *ir* y *venir*, *dar* y *recibir*, *antes* y *después*, *blanco* y *negro*, *bueno* y *malo*, *frío* y *caliente*, etc. (los adjetivos se prestan particularmente a la formación de relaciones antonímicas). A diferencia de las relaciones de sinonimia, las de antonimia manifiestan claras oposiciones de significado, que forman pequeñas estructuras en la lengua. Pese a ello, muchos pares de vocablos que se consideran antónimos tienen significados y usos que contravienen esa idea. Por ejemplo, como vimos en el capítulo 4, el verbo *ir* tiene un valor de auxiliar con el que se forman perífrasis de futuro: "voy a verte mañana", en tanto que *venir* no lo tiene y no se forman las perífrasis opuestas. Son pocos los casos en que la antonimia entre dos signos pueda considerarse fija en la lengua, como en *blanco* y *negro*, *bueno* y *malo*. Por lo general, se producen usos antonímicos según las necesidades del contexto. Muchos "diccionarios de antónimos" abusan de la afinidad de los significados de los vocablos que ponen en oposición, haciendo creer que son estructuras antonímicas fijas.

3.3. La homonimia

A diferencia del fenómeno de la *polisemia*, explicado en el capítulo 4, que manifiesta la naturaleza de la formación del significado de las palabras, el de la *homonimia* es un fenómeno del desarrollo histórico de las palabras, que consiste en una evolución accidental del significante de un signo, que lo lleva a identificarse con otro (*homonimia* quiere decir 'igualdad de nombre'). Por ejemplo: en el acervo léxico del español tenemos dos signos de significante idéntico, que marcaremos como *bolsa*[1]

y *bolsa*[2]. El primero procede del vocablo latino *bursa*, que quiere decir 'objeto de material flexible, generalmente de tela o de cuero, de diferentes formas, que sirve para contener y llevar algo dentro de él'; el segundo, del apellido —*van der Burse*— del fundador flamenco de la primera bolsa de valores en la ciudad de Brujas, en 1409. Aunque el apellido de ese personaje flamenco tiene su origen en la misma palabra latina —lo que complica la historia— y es lo que causó la identidad del significante, los dos signos son diferentes y no corresponden a un fenómeno de polisemia. Tenemos dos *signos diferentes, de significante idéntico*: ambos son *homónimos*.

Estas evoluciones no son sistemáticas. Dan lugar a pequeñas estructuras léxicas debido a los significantes de los signos que entran en relación. Para el hablante muchas veces es incomprensible su existencia, por lo que tiende a buscarles explicaciones de carácter polisémico, que generalmente conducen a la formación de *etimologías populares*, de las que se hablará en el capítulo 11. Así por ejemplo, tratará de explicar *bolsa*[2] diciendo que también en las bolsas de valores 'se guarda dinero'.

Muchas veces, en cambio, un fenómeno originariamente polisémico llega a dar como resultado un alejamiento tal de los varios significados de un vocablo, que no puede reconocerse ya alguna relación entre ellos. Un hablante no sabrá explicar la relación entre un significado y otro; un lingüista, en cambio, prefiere separarlos y ofrecerlos como casos de homonimia. Es lo que sucede muchas veces en los diccionarios, en donde conviene ofrecer varias entradas homonímicas a su lector, para facilitarle la consulta. Por ejemplo: es verdad que el significado de *banco* 'institución de depósito y manejo de dinero' proviene del hecho de que a finales de la Edad Media los primeros "banqueros", que se dedicaban a cambiar moneda en las ferias comerciales, ejercían su trabajo sentados sobre bancos, y que de allí viene el significado actual de esa palabra (marquémosla como *banco*[2]). Sin embargo, esa relación ya no se puede reconocer hoy en día, por lo que conviene distinguir dos homónimos: *banco*[1] y *banco*[2]; si agregamos un *banco*[3], para el significado 'conjunto de peces que nadan juntos', tendremos tres signos homónimos en el español actual.

Podemos concluir que el estudio de la homonimia, en cuanto fenómeno del desarrollo histórico de las palabras, corresponde al estudio de la historia de las lenguas; en cuanto concepto instrumental de la lingüística, corresponde al arsenal de que dispone un lingüista o un lexicógrafo para mejorar la eficacia explicativa y de presentación de los estudios léxicos. Cabe reconocer, sin embargo, que la interpretación rígida de la teoría del signo lingüístico de Saussure, según la cual a un significante sólo puede corresponder un significado, lo que equivale a negar la polisemia, da como resultado una disolución del fenómeno significativo de la polisemia de los vocablos en tantos vocablos homónimos como significados diferentes encontremos.

4. SEMASIOLOGÍA Y ONOMASIOLOGÍA

Hasta aquí hemos venido dándole prioridad al estudio inmanentista de las lenguas, es decir, al estudio de cada lengua en sus propias características. Un análisis semántico del léxico del español, por ejemplo, no puede resultar idéntico a un estudio del léxico del tzeltal o del francés, pues en cada lengua el significado de sus vocablos se ha constituido de manera diferente. Cierto que los análisis léxicos del francés y el español, e incluso del ruso y el húngaro resultarán más parecidos entre sí que los análisis léxicos del español y el tzeltal o el náhuatl; así como es probable que los análisis léxicos del tzeltal, del náhuatl y de alguna de las lenguas zapotecas resulten más parecidos entre sí que con el español, pues esas lenguas forman parte de grandes unidades culturales del mundo, que dejan su huella en sus vocabularios: francés, español, ruso o húngaro forman parte de la gran unidad cultural europea, en tanto que náhuatl, tzeltal y zapoteco forman parte de la gran unidad cultural mesoamericana.

Llamaremos a los análisis inmanentes de la semántica del léxico de cada lengua análisis *semasiológicos*.

Sin embargo, la lingüística ha sido una ciencia que, desde sus orígenes a principios del siglo XIX, ha encontrado en la *comparación* entre lenguas su instrumento de análisis por excelencia. Pues aunque el motivo principal para comparar lenguas en sus primeros años era la búsqueda del parentesco entre diferentes lenguas y el afán de encon-

El término *semasiología* 'estudio del significado' se entendía en el siglo XIX como 'semántica', hasta que este último vino a sustituirlo en la obra del filólogo francés Michel Bréal, *Essai de sémantique (science des significations)*, publicado en 1900. Posteriormente, el término *semasiología* se restringió para significar el estudio del significado de los vocablos de una sola lengua, en sí misma.

trar una lengua originaria —una idea determinada no sólo por la creencia de que el *Génesis* era un libro histórico, sino por el interés de mostrar las relaciones de las lenguas europeas modernas con sus orígenes clásicos en el latín (tratándose de las romances), o con las grandes lenguas de la Antigüedad, como el griego o el sánscrito y, en último análisis, con el supuesto indoeuropeo o indogermánico—, la verdad es que sólo la comparación entre lenguas ofrece al lingüista puntos de apoyo para explorar y penetrar las características de cada lengua; dicho de otra manera, la comparación entre lenguas tiene un determinante valor heurístico.

De ahí que muy pronto se haya desarrollado un método de trabajo diferente del semasiológico: a partir de algún "concepto" considerado como necesariamente universal, o al menos muy general, se buscan las maneras en que cada lengua lo significa; un "concepto" universal o muy general vale entonces como un *tertium comparationis*, un instrumento de comparación del método lingüístico, que sirve para todos los estudios comparativos entre lenguas. Llamaremos a este método *onomasiología*.

Su nacimiento, en la obra del lingüista austriaco Adolf Zauner, *Die romanische Namen der Körperteile. Eine onomasiologische Studie*, Viena, 1902 ("Los nombres romances de las partes del cuerpo. Un estudio onomasiológico"), correspondió a un momento de la historia de la lingüística en que la geografía lingüística estaba comenzando su pleno desarrollo y necesitaba de métodos que ayudaran a poner en relación las diversas evoluciones del léxico, en particular el de las lenguas romances. La onomasiología se convirtió en el método de trabajo de la escuela llamada *Wörter und Sachen* ("Palabras y cosas"), que dio valiosos frutos a la investigación dialectológica e histórica.

Tomemos por caso la investigación de Zauner: le interesaba investigar cómo habían evolucionado los nombres romances de las partes del cuerpo: boca, corazón, rodilla, ceja, nariz, etc. Para hacerlo tomó como *tertium comparationis* los nombres latinos de esas partes, bajo la muy justificada hipótesis de que los nombres franceses, españoles, catalanes, portugueses, italianos, rumanos, retorromanos, etc. habrían evolucionado a partir de ellos. No se trataba de "conceptos" universales, sino de desarrollo histórico de un vocabulario. El método onomasiológico resulta entonces un instrumento muy útil de los estudios etimológicos e históricos, cuando se aplica a familias de lenguas o cuando se le utiliza en geografía lingüística para explicar diferencias entre dialectos. Si se consultan los mapas léxicos del *Atlas lingüístico de México*, del *Atlas lingüístico etnológico de Colombia* o del *Atlas lingüístico etnológico de Andalucía* se podrá ver la manera en que opera.

No es lo mismo un *tertium comparationis*, que es cualquier concepto ajeno a cualquier lengua que, como tal, sirve como principio de la comparación entre lenguas, que una *base de la comparación* que, como hemos visto, es un principio de análisis semasiológico.

Pero de ahí a convertir esas palabras que sirven como *tertium comparationis* en "conceptos" universales hay un pequeño paso, que para muchos lingüistas resulta imperceptible, sobre todo cuando no disponen de una teoría del signo como la saussureana y, en cambio, creen que los signos no son sino meros soportes materiales de una referencia a conceptos ya alojados en la mente o a objetos necesariamente reales para todos los seres humanos. En ese momento "boca" o "corazón", "ceja" o "rodilla" se vuelven conceptos o partes del cuerpo humano que deben tener nombre en todas las lenguas del mundo, con independencia de sus propias maneras de hacer inteligible su experiencia histórica de la vida. Cuando se cree que realmente hay conceptos universales independientes de su formulación en las lenguas del mundo, la investigación lingüística se desvirtúa y aparecen graves errores de análisis. Es imposible y por eso mismo inconveniente tratar de elaborar un catálogo de conceptos universales que sirvan como instrumento del estudio onomasiológico, pues la experiencia de la vida varía de pueblo en pueblo y de lengua en lengua. Si, por ejemplo, en español hay una diferencia clara entre los significados de *ser* y *estar*, al tratar de convertirlos en conceptos universales echaremos en falta su inexistencia en lenguas como el inglés, en la que hay un solo verbo para ambos significados del español: *to be*; si en español *bordo* significa 'orilla ligeramente elevada de un canal', *borde* 'orilla de algún objeto' y *borda* 'orilla de la cubierta de un barco', veremos la polisemia correspondiente del francés *bord* como un "defecto", en vez de reconocer que la comparación entre lenguas es solamente un elemento heurístico para el análisis de cada una de ellas.

Ha habido algunos intentos, fallidos, por construir un sistema de "conceptos" universales para el análisis semasiológico: el primero fue la obra de Rudolf Hallig y Walter von Wartburg, *Begriffssystem als Grundlage für die Lexikographie* (Berlín, 1952); uno más reciente, que desconoce la historia de estos esfuerzos onomasiológicos, fue el de Anna Wierzbicka, *Lingua mentalis: The Semantics of Natural Language* (Sydney, 1980).

4.1. El cálculo onomasiológico

Pero señalar las concepciones equivocadas y los trabajos fallidos no impide ni niega el valor de los planteamientos onomasiológicos, siempre que se proceda de la manera siguiente:

a] No tomar como "conceptos" vocablos de ninguna lengua particular, sino construir los conceptos sobre la base de oposiciones binarias, formuladas mediante perífrasis muy detalladas en la lengua que sirva como lenguaje de descripción, o mediante reducciones esquemáticas de cierta clase de relaciones.

b] No pretender cubrir todo el léxico de una lengua con una sola sistemática, que lleve a la formación de verdaderas "pirámides conceptuales", sino construir relaciones conceptuales parciales para campos léxicos cuya cantidad de vocablos sea cerrada.

Llamaremos *noemas* a los conceptos resultantes de tales construcciones onomasiológicas. Veamos el siguiente ejemplo:

Para analizar las palabras (o los morfemas, en muchas lenguas) que significan la localización de una persona o de un objeto en relación con una acción verbal se puede distinguir al 'participante en la acción verbal, que habla y ejecuta la acción', en relación con el cual se define el punto de origen de la acción, su "punto cero" (se puede representar con AØ); a él se opone binariamente 'quien participa en la acción verbal, que escucha', representado con A~Ø. Podemos suponer que esta distinción sea universalmente válida para analizar esa clase de *déicticos* (el vocablo *déictico*, del griego *deixis*, quiere decir 'indicador'); es decir, que puede sernos útil para analizar cualquier lengua. A estos dos participantes en la acción verbal, podemos oponer 'quien no participa en la acción verbal y no la presencia, es decir, no está en relación con el hablante': ~A~Ø.

Si exploramos la manera de significar estos tres conceptos, producto de un cálculo onomasiológico, vemos que, en español, tenemos:

'AØ' *yo* / 'A~Ø' *tú* // '~A~Ø' no-yo, no-tú

(Si se combinaran mecánicamente los conceptos construidos, tendríamos también ~AØ, que resulta una contradicción, pues es imposible no ejercer la acción verbal y ser su punto de origen.)

Para continuar analizando el noema '~A~Ø' podemos introducir ahora todos aquellos puntos o entidades que, sin participar en la acción verbal, pueden referirse a los dos participantes: '~A~Ø (Ø)'/ '~A~Ø (~Ø)'. Con base en estos cuatro noemas encontramos, en francés:

'AØ' *moi* / 'A~Ø' *toi* / '~A~Ø (Ø)' *celui-ci* / '~A~Ø (~Ø)' *celui-là*

Y podemos ampliarlo a tomar en cuenta su referencia o no a la acción verbal: '~A~Ø (~ØA)'/ '~A~Ø (~Ø~A)', lo que nos explica, en español:

'AØ' *yo* / 'A~Ø' *tú* / '~A~Ø (Ø)' *éste* / '~A~Ø (~ØA)' *ése* /
'~A~Ø (~Ø~A)' *aquél*

Sobre esa base, se puede introducir la distinción entre una deixis local o mostrativa, y una deixis personal, es decir, una indicación de un punto o una indicación de una persona; en ese caso tendríamos, en francés:

'AØ' *moi* / 'A~Ø' *toi* / '~A~Ø (Ø)' *lui*

mientras que en español:

'AØ' *yo* / 'A~Ø' *tú* / '~A~Ø (Ø)' *él*

Podemos ahora introducir las diferencias de sexo, simbolizadas con Pm (persona masculina)/Pf (persona femenina): en árabe, 'A~ØPm' se significa con *anta*, y 'A~ØPf' con *anti*. Por el contrario, es probable que no muchas lenguas identifiquen el sexo del hablante 'AØ', debido al hecho de que el hablante está presente a la vista de quienes participan en la acción verbal, o que su timbre de voz lo identifique. Igualmente, hay muchas lenguas que no hacen distinción entre persona gramatical y deixis locativa. Pensemos que en español podemos escuchar quien diga: "¿quién hizo esta travesura? —*Éste*", en vez de *él*. Hay lenguas en donde no se identifica directamente el sexo de los participantes en la acción verbal, sino la igualdad o la diferencia de sexo entre quien habla y a quien refiere, es decir: 'la persona del mismo sexo que quien habla' frente a 'la persona de sexo diferente de quien habla'.

Como se puede ver en este pequeño ejemplo, la onomasiología no supone la existencia real de "conceptos" anteriores a toda lengua, y que sean necesarios para todos los seres humanos, una actitud propiamente platónica. Por el contrario, la onomasiología construye los conceptos que requiere un análisis lingüístico, orientados por la clase de fenómenos semánticos que se desea analizar, y como un juego de relaciones opositivas de carácter binario, cuya función es contribuir a comparar los significados de los signos lingüísticos de diferentes lenguas e iluminar mejor las características particulares del significado de los signos de cada lengua.

Este pequeño ejemplo da una idea del método onomasiológico. Se puede seguir para comparar diferentes paradigmas cerrados de las lenguas, como el tiempo y el aspecto verbal, así como paradigmas de clasificadores semánticos en lenguas como la maya o la mixteca. Pero hay que insistir en que la naturaleza de los noemas no proviene de unos conceptos universales, pertenecientes a las ideas platónicas, sino de una construcción analítica del lingüista.

5. CONCLUSIÓN

En este capítulo hemos introducido los elementos básicos del análisis cualitativo del significado de los vocablos, así como los principales razonamientos que sirven para plantearse un análisis semántico.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

base de la comparación; valor de campo; rasgo significativo; sema; campo asociativo; estado de lengua; sincronía; unidad dialectal; comunidad de comunicación; tradición verbal culta y popular; significado y designación; análisis binario; semema; microestructura léxica; semasiología; onomasiología; noema; sinonimia; antonimia; homonimia; hiperonimia.

OBRAS CONSULTADAS

BÜHLER, Karl, *Teoría del lenguaje*, Madrid, Revista de Occidente, 1967 (1a. ed. alemana, Jena, 1934).

COSERIU, Eugenio, *Principios de semántica estructural*, Madrid, Gredos, 1981.

HEGER, Klaus, "Deixis personal y persona gramatical", en *Teoría semántica, Hacia una semántica moderna*, II, Madrid, Alcalá, 1974.

MELCHUK, Igor A., *Dependency Syntax. Theory and Practice*, Albany, SUNY, 1988.

MOUNIN, Georges, "Essai sur la structuration du lexique de l'habitation", *Cahiers de Lexicologie*, 6, I (1965), pp. 9-24.

10
EL LÉXICO, SÍMBOLO SOCIAL

INTRODUCCIÓN

Como hemos venido viendo, la unidad *palabra* tiene un papel privilegiado en la percepción y la reflexión social acerca de las lenguas, debido a su característica central de nombrar objetos, acciones y relaciones. Precisamente por esa característica y ese papel, las palabras se convierten fácilmente en *símbolos sociales*; es decir, *trascienden su naturaleza de signos lingüísticos y se convierten en representantes de concepciones, valores y tabúes sociales*, a los que se les atribuyen desde propiedades mágicas hasta funciones morales o ideológicas. En este capítulo nos ocuparemos de explicar estos fenómenos y relacionarlos con sus efectos normativos sobre el léxico, que tienen un importante papel en la manera en que las sociedades consideran sus lenguas.

1. CONCEPTOS BÁSICOS

Durante los primeros años de vida de los seres humanos la experiencia de la vida va *indisociablemente* acompañada, en muchos momentos, de los signos lingüísticos con que la significamos. Como vimos en la primera parte, cuando un niño aprende la palabra *mamá*, ésta está tan imbricada con todas las experiencias que rodean el contacto con su madre, que no sólo llega a pronunciar la palabra cuando ve a su madre o a un objeto que suele acompañarla, sino que incluso llega a esperar que su mera invocación mediante la palabra tenga un efecto práctico inmediato, haciéndola aparecer (acerca de esto tratamos en el capítulo 2, § 2.1). Entre el signo *mamá* y aquello a lo que refiere, la madre, se crea una relación de *participación*, por la cual el niño cree que la palabra *mamá* forma parte del ser al que se nombra. Este fenómeno se encuentra entre las concepciones más antiguas de

muchos pueblos de la Tierra a propósito de la relación entre las palabras y las cosas y contribuye a la formación de creencias mágicas, que son la base de muchos encantamientos, hechizos e invocaciones supersticiosas.

1.1. El tabú verbal

Esa clase de creencias da lugar al *tabú* verbal. El fenómeno social y psicológico del *tabú* consiste en creer que, porque hay relaciones de participación entre la naturaleza de ciertos animales, plantas, objetos e incluso personas y miembros de la sociedad o grupos de ella, y los vocablos que los significan, no deben pronunciarse, so pena de recibir castigos o daños quien lo haga; si las palabras participan de la naturaleza de aquello que significan, entonces tienen poder mágico: si se pronuncia la palabra *diablo*, su efecto inmediato es la invocación al Diablo; si para el pueblo judío no se debe invocar el nombre de Dios en vano, había que encontrar otra manera de hablar de él: "El que es", *Yahvé*. En español, por ejemplo, hay muchas maneras de nombrar al diablo, remplazando *diablo* por otras palabras; en México, *pingo*, *pingorongo*, *patas de chivo*, *diantre*, *dianche*, *el maligno*, *el enemigo*, *chamuco*, *cachudo*, *pituche* e incluso *sanquisinito* (según se puede comprobar en el *Atlas lingüístico de México*); en España, *ángel del mal*, *ángel caído*, *cornudo*, *diantre*, *diaño*, *chápiro*, *rabudo* y muchas más. Como la muerte se suele personificar del mismo modo que al diablo, las tradiciones verbales populares en México han creado multitud de *eufemismos*: *la calaca*, *la parca*, *la pelona*, *la pelada*, *la dientuda*, *la flaca*, *la huesuda*, *la tilica*, *patas de catre*, *la dama de la guadaña*, *la catrina* (a partir de los grabados de José Guadalupe Posada, a comienzos del siglo XX) y, por supuesto, *la chingada*, con su posterior eufemismo de *la tiznada*, ya que aquella palabra también se tiende a eludir socialmente, como lo trataremos más adelante.

Quizá a esa clase de fenómeno se deba también la cantidad de nombres que ha recibido la *mustela* (su nombre latino), con que se nombraba a la *Mustela frenata* (su nombre técnico zoológico) en todas las lenguas europeas en las que se sustituyó por aparentes eufemismos: en España, en territorio castellano, *comadreja*; *mostolilla* en el norte de Castilla la Vieja; *donecilla* y *donociña* en Galicia; *mustela* en Valencia y Cataluña, y *paniquesa* en Aragón (más otros nombres); *belette* 'mujercita bella' en francés; *donnola* 'damisela' en italiano, y de maneras similares en varias lenguas europeas. Según relata Sebastián de Covarrubias, autor del primer diccionario de la lengua española, el *Tesoro de la lengua castellana o española* (1611), la comadreja se ocupa de limpiar toda clase de sabandijas en el interior de las casas, por lo que se la compara con el ama de casa (*comadreja*, *donecilla*, *donnola*). En náhuatl, al *tlacuache*, una zarigüeya (marsupial) que abunda en Mesoamérica, objeto de complejas relaciones míticas, se le nombra también con eufemismos: *ilamaton* 'viejecita', *cihuatlamacazque* 'sacerdotisa' y *maxtle* 'taparrabo'.

La palabra *tabú*, tomada probablemente de una lengua del archipiélago de Tonga, en la que quería decir 'prohibido', e introducida primero al inglés por el capitán Cook en 1785, en su relato de viajes por Oceanía, se ha convertido en un término técnico de la etnología. Hay una nutrida bibliografía respecto al tabú, además de *La rama dorada. Magia y religión*, de James George Frazer, *Tótem y tabú* de Sigmund Freud o *El pensamiento salvaje* del antropólogo francés Claude Lévi-Strauss.

Llamamos *eufemismo* a todo vocablo utilizado para eludir la pronunciación o la escritura de otro y sustituirlo en una acción verbal. Su raíz griega, εὐφημός, quiere decir 'que habla bien o que no usa palabras de mal agüero'; nótese la relación etimológica del vocablo con la magia religiosa.

Son motivos de la formación de eufemismos, además del tabú religioso, también las personas, los animales y los objetos que, por diferentes motivos, se consideran "impuros", como los excrementos humanos, los órganos sexuales o, en general todo lo que tenga que ver con la sexualidad: *caca*, *popó* en el vocabulario que usan los adultos con los niños, *hacer del dos*; *panocha*, *papaya* y *mono* para la vagina; *pito*, *pirulí*, *pinga*, *pilín*, *picha*, *consolador*, *palo*, *chile* para el pene; *busto*, *seno*, *chichi* (aunque es su nombre náhuatl) y hoy en día, por influencia angloamericana, *boobi*, etc. para las mamas femeninas; *trasero*, *pompis*, *nailon* para las nalgas.

1.2. La motivación de los nombres

En el famoso diálogo *Cratilo* de Platón, la cuestión que da origen a la conversación entre Sócrates y sus interlocutores Hermógenes y Cratilo es si "cada cosa tiene un nombre que le es *naturalmente* propio", lo cual

no indica necesariamente que ahí se discutiera la *participación* del signo en la naturaleza de la cosa, como en el caso de las creencias mágicas, sino su *motivación*. Platón desarrolla tres argumentos: el de la absoluta convencionalidad de los nombres de las cosas, que pueden variar a voluntad de cada individuo (recordemos el fragmento de *A través del espejo*, citado en el capítulo 2); el de una "arbitrariedad justificada" de los "legisladores" de los nombres (¡cuántos académicos de la lengua se adherirían con gusto a ese argumento!), y el de la *motivación* del nombre en la esencia de la cosa. Este último argumento propone que el origen de los nombres se encuentra motivado en una *comprensión verdadera de la esencia de la cosa nombrada*; no forma parte de ella, sino que obedece a la comprensión de la cosa, de quienes hayan inventado la palabra. Así por ejemplo, Sócrates sostiene que el nombre de *Orestes*, el hijo de Agamenón y Clitemnestra (léase la famosa trilogía *La Orestíada*, de Eurípides), recibió su nombre motivado por su naturaleza "bravía, salvaje y montaraz", que es el significado de la palabra griega ὀρεινόν.

En la enseñanza de la semántica léxica, las diferencias entre las dos primeras concepciones (arbitrariedad-convencionalidad absoluta y relativa a la decisión de los "legisladores") se reúnen en una sola, de la *arbitrariedad-convencionalidad*, que se enfrenta con la de la *motivación*, y se caracterizan como "doctrina de la *thesei*", la primera, y "doctrina de la *physei*", o motivación, la segunda. La primera se atribuye, simplificándolo, a Aristóteles. Ningún lingüista contemporáneo sostendría hoy en día la segunda, pero la concepción sigue manifestándose en muchos hablantes de todas las lenguas y tiene su importancia en la investigación etimológica, como se explicará en el capítulo 11. En cuanto a la doctrina de la *thesei*, aunque es claro que los signos no tienen motivación y mucho menos participación en la naturaleza de lo que nombran, tampoco son producto de una convención, sino de una *tradición verbal* en cada lengua histórica.

2. EL EFECTO NORMATIVO DEL SÍMBOLO SOCIAL

Si el fenómeno del tabú nos permite comprender con facilidad el motivo de la formación de eufemismos y darnos cuenta de que con ellos la palabra trasciende su carácter de signo lingüístico para volverse símbolo social, podemos ahora considerar cómo, de manera más general, otros valores de la sociedad toman como objeto las palabras y dan lugar a reglas para su uso, cuidadosamente transmitidas a los miembros de la comunidad lingüística desde muy temprana edad. Si, por ejemplo, exclamar ¡*carajo*! se juzga incorrecto en un habla respetuosa, hay una regla social que propone sustituir la expresión por un eufemismo, emparentado en su forma de la expresión, como ¡*caray*! o ¡*caramba*!, socialmente aceptados o al menos tolerados; si decir *chingar* resulta ofensivo para muchas personas, en la tradición popular mexicana se ofrece *tiznar* como eufemismo, o incluso *chiflar* ("mandar a alguien a chingar a su madre" se convierte en "mandarlo a chiflar a su *máuser*"; máuser es el nombre de un rifle). En Cuba se considera un vocablo grosero *papaya*, por ser un eufemismo de la vagina, y se le sustituye por otro eufemismo, socialmente aceptado: *fruta bomba*; *culo* en México es grosero, y por eso se le sustituye por sus eufemismos, incluso por la expresión "en donde la espalda pierde su honesto nombre"; hay quienes prefieren no nombrar a los *huevos* (por ser ésta eufemismo de los testículos) y llamarlos *blanquillos*; un adolescente actual preferirá no decir frente a sus amigos que algo es *hermoso*, y decir en México que *está chido*, o en España, que *es guay*, en tanto que un adulto que utilice estas expresiones en un medio académico o en una ceremonia formal, quizá dé lugar a burlas o a verdadera molestia; no se puede llamar a un policía en México *azul* (por el color del uniforme) o *mordelón* (por ser *mordida* el nombre popular del soborno); resulta ofensivo para un natural de España llamarlo en México *gachupín*, o a un natural de la Ciudad de México, *chilango*. Ya vemos en inglés norteamericano los eufemismos creados para que las palabras *negro* o *black* no sean manifestación de discriminación racial: *color people*, *african-american person*, etc., junto con sus correspondientes *disfemismos* —es decir, vocablos utilizados para manifestar cruda y agresivamente un significado: *neger*; en México, *indígena* como eufemismo de *indio*, e *indio patarrajada* como disfemismo. Los políticos de todo el mundo suelen buscar afanosamente eufemismos para disfrazar conceptos que les parecen poco atractivos socialmente o que podrían acarrearles protestas sociales, como llamar "daños colaterales" a toda la población civil que matan las acciones militares en Palestina, en Afganistán o en Irak; a los viejos, "personas de la tercera edad" o, peor, "adultos en plenitud" por el gobierno de Vicente Fox en México.

Todos los *eufemismos* anteriores y los nombres populares de muchos objetos o acciones responden a la acción de reglas sociales determinadas por la cortesía, por una concepción de la moralidad, por la necesaria solidaridad de los grupos juveniles, por las diferencias de generación entre las personas, por el respeto a los seres humanos, o a reglas que imponen los gobiernos a sus administraciones para impedir que haya protestas sociales. Tales reglas son *normas sociales*, es decir, reglas que determinan lo que es correcto en situaciones determinadas del contacto social.

Es característica del léxico convertirse fácilmente en objeto de esa clase de normas sociales, que *no hay que confundir con las normas lingüísticas* (aunque las normas lingüísticas son parte de las sociales). Pues las normas lingüísticas no trascienden a los signos lingüísticos; no los convierten en símbolos sociales, sino que actúan sólo para conservar ciertos elementos de la lengua histórica que han sido históricamente valorados por su cultura de la lengua y sus necesidades de eficacia y precisión. Así por ejemplo, las reglas de ortografía, como vimos en el capítulo 5, son normas lingüísticas cuyo objetivo es regular algunos elementos del sistema de escritura, para hacerlo más económico y eficaz. Las normas sintácticas, como las del uso del gerundio en español, por ejemplo: *no se debe decir*: "una caja conteniendo 12 botellas" *sino* "una caja que contiene 12 botellas", o las que determinan las preposiciones regidas por ciertos verbos, como *subir a* en vez de *subir en*, *esperar a* alguien y no *esperar por* alguien (una influencia del inglés), son reglas sintetizadas de la experiencia gramatical de la cultura de la lengua, no son reglas de uso de símbolos sociales.

Aunque la escritura da lugar a lo que el filólogo argentino-venezolano Ángel Rosenblat llamaba "fetichismo de la letra", que es un caso claro de trascendencia de la letra a símbolo social. Pensemos en el prestigio que tiene, para los mexicanos, la <x> de *México*, que lleva a algunos individuos a modificar su nombre: *Xorge*, *Xuan*; las grafías inglesas <ss> y <cc> llevan a ciertas personas a llamar a sus hijas *Vanessa*, *Issabel*, *Rebecca*, y se toman como prestigioso extranjerismo; incluso hay quienes escriben *lasser*, aun cuando el acrónimo inglés sea *laser* y, en español, *láser*. Quizá un caso extremo de fetichismo es la marca de helados *Häagen-Dazs*, que no corresponde a ningún sistema de escritura, aunque parezca germánico. Por el mismo fetichismo un sastre opta por llamar a su sastrería *D'Nino* (en las escrituras que abrevian la preposición *de*, como el francés o el italiano, esto sólo puede hacerse cuando el fonema de la palabra siguiente es vocálico: *d'amour*, *d'Este*).

Al no ser las normas sociales que se aplican al léxico de la misma clase que las normas lingüísticas que se aplican al resto del sistema de la lengua, especialmente a la ortografía y la sintaxis, su validez depende de cada comunidad lingüística, varía más fácilmente con el tiempo y opera sobre vocablos seleccionados, no sobre parcelas completas del léxico. Por ejemplo, Don Quijote recomendaba a Sancho no usar la palabra *regoldar*, por ser poco fina socialmente y, en cambio, la corregía por *eructar*; hoy en día, *eructar* se sustituye por el eufemismo *repetir*.

2.1. El disfemismo

El hecho de que las normas que se ponen en práctica en el léxico sean sociales y obedezcan a valores determinados por la sociedad, como la cortesía, la urbanidad o la moralidad, da lugar al fenómeno contrario: el gusto de ciertos grupos sociales o de algunos individuos por contravenir esas normas, como manifestación de desdén o de desprecio por los valores socialmente instituidos. El disfemismo cumple esa función. En una situación revolucionaria, en la que se ha logrado romper un orden instituido y tiende a imponerse otro, como sucedió en México después de la Revolución de 1910, en Cuba después del derrocamiento de Batista y la instauración de la Revolución cubana, o en Nicaragua, con la llegada al poder del sandinismo, varios grupos sociales optan por un lenguaje disfemístico, en el que las "palabras fuertes" predominan. Lo mismo sucede con los adolescentes, cuyo proceso de maduración requiere un distanciamiento claro de la generación de sus padres. Así, es común que los jóvenes tomen palabras de léxicos de grupos socialmente marginados o incluso de la delincuencia. En el español de México es lo que ha sucedido con vocablos como *chido*, que quiere decir aproximadamente 'bueno, hermoso, divertido', *afanarse* 'trabajar', *rolar* 'circular', *apañar* 'robar o conseguir algo', *baros* 'pesos (moneda)', todos ellos procedentes del caló nacional (de lo que se hablará en seguida).

2.2. Los léxicos de solidaridad

Esta clase de grupos sociales, como los jóvenes o los delincuentes "de oficio", como los carteristas, los ladrones de casas, los comerciantes de drogas al menudeo (sobre todo de mariguana, que ha sido tradicional en México), tienden siempre a separarse de los valores instituidos en su comunidad y, por ello, a hacerse de un vocabulario propio, que a la vez

que los distinga del vocabulario común, les sirva como seña de identidad con otros miembros del grupo. Se forman así las *germanías*, *argots*, *jergas* o *caló*.

Germanía es el nombre histórico en español de los grupos de delincuentes, especialmente ladrones y contrabandistas, que desde muy antiguo se constituían como hermandades (la palabra *germanía*, tomada del catalán y ésta del latín *germanus* 'hermano carnal' —que dio *hermano* en español—, se convirtió en nombre de las hermandades de delincuentes que desde el siglo XVI en Valencia o en Sevilla se hicieron notar en esas sociedades). Tales hermandades de delincuentes se comportaron como sociedades secretas en Sevilla hasta el siglo XVIII, es decir, de manera semejante a las *mafias* sicilianas, la *camorra* o la *cosa nostra* napolitanas. Su vocabulario, diferente del del resto de la sociedad, que sólo se interesaban por aprender los delincuentes, se comenzó a considerar secreto, como una especie de código cifrado, que impidiera su comprensión por parte de la policía. Hasta hoy en día se piensa que la germanía es un lenguaje secreto, que cambia cada vez que alguien logra descifrarlo. La realidad es otra: el léxico germanesco es un vocabulario muy tradicional, cuya comprensión no se impide a nadie, al grado de que sus estudiosos suelen formar parte, precisamente, de la policía. Se puede demostrar que las germanías hispanoamericanas, como el *caló* mexicano, la *replana* colombiana o incluso parte del *lunfardo* argentino (porque lo forman varios cientos de palabras de origen italiano), provienen históricamente de la germanía sevillana del siglo XVIII, llamada comúnmente "nueva germanía", para distinguirla de la "vieja", de la que se encuentran buenos ejemplos en las novelas picarescas españolas, como el *Lazarillo de Tormes* o *La vida del Buscón*, de Francisco de Quevedo.

Tal "nueva germanía" recibió aportaciones de la lengua de los gitanos (procedentes de India, a través del norte de África) llamada *caló*, desde su llegada a la península ibérica a finales del siglo XVI. Desde entonces, y sobre todo en México, se considera que *caló* es el vocabulario de los delincuentes (en especial, ladrones o rateros, y comerciantes de drogas; no los grandes narcotraficantes, que suelen ser políticos y financieros respetados). En México y el suroeste de Estados Unidos de América, el *pachuco* es una variedad del caló mexicano.

Aunque también, antes de que se generalizara el vocablo *caló*, era común llamar a la germanía *jerigonza*, tomada del francés *jargon*, que también dio *jerga* en español. Estas últimas, así como el vocablo francés *argot*, pasaron a nombrar cualquier vocabulario especializado de diversos gremios, que no llegue a constituir un vocabulario técnico o científico.

En todos los casos se trata de vocabularios de solidaridad entre los miembros de un gremio, que les permite reconocerse y entablar diálogos muy eficaces. Hay jerga de médicos en los hospitales, de ingenieros, de albañiles, de mecánicos de automóviles, de técnicos de computación, de pilotos de avión, de estudiantes, etc. De la jerga de estos últimos podemos recordar *tronar* 'reprobar un examen', *acordeón* —llamado en España *chuleta*— 'pequeño resumen del contenido de una materia, que se saca a escondidas durante un examen', *barco* 'maestro que no es exigente', *pasar de panzazo* 'aprobar con la mínima calificación', etcétera.

Todas las jergas, germanías o calós son básicamente vocabularios, constituidos por sustantivos, verbos y algunos adjetivos y adverbios. Difícilmente se encuentran en ellos elementos sintácticos propios, que puedan hacer pensar que se trata de lenguas completas. En caló mexicano, una expresión como *apañar jando baril*, que quiere decir 'robar mucho dinero' es de las pocas que van más allá de los vocablos aislados.

2.3. Las groserías

Hay vocablos que tienen una función social muy diferente de la de los eufemismos, pues buscan insultar, provocar reacciones de molestia o de enojo en otras personas. A diferencia de un hablar ofensivo, que puede consistir de palabras cultas e incluso elegantes, la grosería llama la atención en sí misma. Puede ofender a cualquier persona que la escuche, aunque no esté dirigida a ella. Por ese carácter, las sociedades las proscriben. Sin embargo, la violencia que producen cuando se manifiestan, que a veces lleva a los golpes, puede llegar a tener un efecto contrario, apaciguador de ciertas personas, cuando están poseídas por la histeria. Una grosería bien escogida en cierto momento y en cierta situación, aplaca el ánimo y resuelve los conflictos. Incluso, por su agresividad pueden llegar a cautivar a una persona y, en vez de alejarla de quien las pronuncia, acercarla a ella. Tomemos por ejemplo el siguiente piropo a una mujer que caminaba, con paso airoso y gallardo, por una calle de la Ciudad de México: "¡guapa, eso se llama *pisar* y no *chingar* el pavimento!"

2.4. El extranjerismo

Todas las lenguas del mundo son mestizas. Todos los pueblos y las culturas que han habitado la Tierra (¿habrá alguna excepción?) han tenido contactos entre ellos; han aprovechado conocimientos y costumbres de los pueblos vecinos, o de los que los invaden y sojuzgan, como para haber adaptado formas lingüísticas, signos, significados de sus lenguas a las propias. El inglés moderno no sería como es, si no hubiera adaptado una enorme cantidad de vocablos del francés, correspondientes a la época en que los reyes normandos, del norte de Francia, dominaban Inglaterra; el náhuatl aportó signos a las lenguas mesoamericanas con las que tuvo contacto, y dejó regados por todo el territorio mesoamericano nombres de pueblos, de lugares y de lenguas, como *tarasco* para el purépecha de Michoacán y parte de Guanajuato y Jalisco, u *otomí* para el hñahñú de Hidalgo y Querétaro. Al español, el aporte civilizatorio de Al Andalus —la España musulmana— le dejó vocabulario de la hidráulica, de la astronomía, de las matemáticas, de la albañilería, de la organización política y de muchas áreas más de la vida civilizada. Buena parte del vocabulario científico del español entre los siglos XVIII y XIX provino del francés; hoy, del inglés.

Las características del léxico son las que permiten tales adopciones, que resultan más difíciles cuando se trata de palabras gramaticales (a pesar de ello, al árabe le debemos la preposición *hasta* del español; el náhuatl moderno ha adaptado desde la numeración y los días de la semana hasta varias preposiciones del español, por no hablar del vocabulario ganadero y de muchas otras áreas). El inglés actual, lengua de la llamada "república imperial", invade con vocablos a todas las lenguas del mundo.

Precisamente por esos contactos, que son naturales en la vida de las sociedades, la capacidad que tiene el léxico para convertirse en símbolo social da lugar a reacciones en contra del extranjerismo, pues se ve en él un afán de dominación por parte de la sociedad aportadora —concebida como invasora— o de claudicación en el aprecio de la lengua propia, por parte de la sociedad receptora.

El extranjerismo puede consistir de *préstamos*, caracterizados por la conservación del vocablo con todas las características de escritura y morfológicas de su lengua de proveniencia (no del todo con su fonética),

como *hardware*, *software*, *log in*, *timing*, *lobbying*, *check-in*, *cool*, *jogging*, *stand-by*, *jet*, *switch*, *blazer* del inglés actual; *restaurant*, *chiffon*, *toilette*, *chef*, *haute cuisine*, *pret-à-porter*, *brasserie*, *hors d'oeuvre*, *vol au vent*, *maître à penser*, *maître d'hôtel* del francés; *kindergarten*, *zeitgeist*, *weltanschauung* del alemán; *chiao*, *pizza*, *lasagna*, *spaghetti*, *chianti*, del italiano, etcétera.

Puede también consistir de *calcos*, en que se toma el sentido del vocablo extranjero y se le busca una palabra con sustancia de contenido semejante en la lengua receptora, como *cabildeo* para *lobbying*, *jardín de niños* para *kindergarten*, *platillo volador* para *flying saucer*, *tormenta de ideas* para *brainstorm*, *trampantojo* para *trompe-l'oeil*, *alineación* para *line-up*, *jardinero* para *fielder* y *parador en corto* para *short stop* (en beisbol), *cazacerebros* para *brain hunter*, *tenis de mesa* para *ping pong*, *baloncesto* para *basket ball* y *balón mano* para *handball*, *tiempo fuera* para *time out*, etcétera.

Por último, el extranjerismo entra a la lengua receptora como *adaptación* a sus características de escritura, fonológicas y morfológicas, por ejemplo en: *restorán*, *volován*, *rin* (del inglés *reim*, la rueda de los coches), *cabús* (del inglés *cab ooze*), *riel* o *raíl* (del inglés *rail*), *champú*, *futbol* o *fútbol*, *básketbol*, *boxeador* para *boxer*, *piyama* para *pijama*, *suéter* para *sweater*, *céspol* para *cess-pool*, *troca* para *truck* y *yonque* para *junk yard* (en el norte de México), *chofer* para *chauffeur*, *garaje* para *garage*, *mofle* para *muffler* (silenciador de los coches), etcétera.

2.5. El neologismo

Se entiende por *neologismo* todo vocablo de reciente difusión en la vida social. Vocablos técnicos como *chip*, *byte*, *internet*, *holding*; vocablos creados para dar cuenta de nuevas experiencias sociales, como *jibarizar*, que se empezó a utilizar en Chile cuando las políticas económicas neoliberales comenzaron a destruir industrias nacionales; *colombianizar*, difundido por el periodismo para caracterizar el proceso de disolución del tejido social de un país como resultado del crecimiento de las guerrillas y de las bandas de narcotraficantes; *narcodólar*, con que se habla del dinero que ganan los narcotraficantes, etc. son vocablos que se comenzaron a difundir en español hace menos de 30 años. Todos ellos obedecen a la necesidad de significar nuevas experiencias de la vida social. Algunos permanecen y entran a formar parte del léxico de la lengua histórі-

ca, en tanto que otros caen pronto en desuso y desaparecen en poco tiempo.

La formación de neologismos se ha vuelto un campo de investigación y de creación léxica muy importante de las lenguas modernas, pues todas se ven en la necesidad de introducir nuevos vocablos a su léxico, para poder ir al paso, sobre todo, de nuevas invenciones tecnológicas o de nuevos descubrimientos en la naturaleza. La disciplina que estudia la formación de neologismos forma parte de la morfología y de la investigación de la formación de palabras. Se llama *neología*. Pero a la neología también corresponde la creación calculada de nuevos vocablos, por parte de diversas agencias de actualización de las lenguas contemporáneas. Incluso lenguas como el latín del Vaticano o el hebreo moderno, cuyos acervos léxicos provienen de la antigüedad, pero requieren nuevo vocabulario para poder significar al mundo contemporáneo, son campo de trabajo de grupos especializados en neología. Cuando las lenguas amerindias reciban el reconocimiento que siempre se les ha negado y obtengan el apoyo necesario para actualizarlas, requerirán también procesos neológicos.

2.6. El solecismo y el barbarismo

El hecho de que las lenguas históricas evolucionen con cierta regularidad, pero no con absoluta sistematicidad, hace aparecer constantemente en sus léxicos vocablos que se apartan de sus reglas morfológicas y sintácticas; aquellos vocablos que se forman contraviniendo esas reglas, que son internas de la lengua, no impuestas socialmente, se denominan *solecismos*. (El vocablo proviene del griego *Soles*, una colonia griega en Cilicia, a cuyos habitantes acusaban de no saber hablar el griego correctamente.) Son ejemplos de solecismos actuales: *aplicar* a un empleo, en vez de *hacer una solicitud*, y *no aplicar* la cláusula de un contrato o una regla determinada, en vez de *no aplicarse* (el verbo *aplicar* en español es transitivo); construcciones como "fue entonces *que* salió de viaje", "es así *que* logró hacerlo", "será en ese pueblo *que* nos encontremos" en vez de "fue entonces *cuando* salió de viaje", "es así *como* logró hacerlo" y "será en ese pueblo *donde* nos encontremos". El solecismo refleja por lo general falta de conocimiento de las tradiciones verbales que crean la cultura de la lengua; muchas veces obedece a la importación acrítica de palabras o construcciones de otra lengua, como sucede ahora a causa del contacto con la lengua inglesa.

El *barbarismo*, por su parte, caracteriza a cualquier construcción morfológica o sintáctica, así como a cualquier introducción de un nuevo significado de una palabra, que contravienen las tradiciones cultas de una lengua, aunque aprovechen sus características morfológicas y sintácticas, como sucede con *ofertar* algo, en vez de *ofrecer* algo; *honrar* un compromiso, en vez de *respetarlo*; *recepcionar* una pelota en vez de *recibirla*; *explosionar* una bomba en vez de *hacerla explotar*; etcétera.

3. EFECTOS IDEOLÓGICOS SOBRE EL LÉXICO

Vocabularios de solidaridad —germanías, caló, jergas—; palabras tomadas de lenguas extranjeras; vocablos que nombran objetos o significan acciones rechazadas o sancionadas como ofensivas por la sociedad o algún grupo social determinado; solecismos y barbarismos, se convierten fácilmente en símbolos sociales, que se prestan para articular actitudes normativas o prescriptivas de carácter ideológico. Si el disfemismo y la grosería se reivindican como parte de un hablar revolucionario o contestatario, el eufemismo muchas veces manifiesta una moralidad o una urbanidad socialmente compartidas, que en ciertos momentos históricos llega a imponerse sobre toda la sociedad. En francés, por ejemplo, en el siglo XVII, se consideraba ofensivo hablar de la espalda, el pecho o los senos; se proscribieron las palabras *poitrine* y *dos*, que los nombraban, y para eludir hablar de los senos, se impuso *gorge*, que quiere decir 'garganta'. Los diccionarios de la Academia Española tardaron muchos años en dar entrada a palabras de esa clase, como lo muestra el *Diccionario secreto* de Camilo José Cela. Algo semejante, no del orden moral, sino político-racista, sucedió en Alemania durante el nazismo, en donde los movimientos nacionalistas reaccionarios trataron de imponer una germanización del vocabulario alemán, para eliminar de la lengua los vocablos que no tuvieran raíz germánica. Así, palabras de origen griego como *kino* 'cine', o *telephon* se sustituyeron por *lichtspiel* y *fernsprecher*; *hotel*, del francés, por *gasthaus*; *hochschule* por *universität*; otras no perduraron, como *hauptwort* e incluso *dingwort* por *substantiv*; *vernunfttum*

por *rationalismus*; *geistgruppwissenschaft* por *soziologie*; *erdmittigkeit* por *geozentrizität*, etc. (Lo paradójico de este movimiento fue que el Führer se negó a apoyarlo y, por el contrario, introdujo muchos vocablos prestados de otras lenguas. Así, los filólogos y lingüistas racistas le proponían, para 'campo de concentración', *sammellager*, *zwangslager* o *straflager*, pero el partido nazi optó por *konzentrationslager*; en vez de *entmannung* o *unfruchtbarmachung*, optó por *sterilisation*.)

3.1. El purismo

Hay dos ideologías lingüísticas que conviene considerar en relación con el valor simbólico de los vocablos: el *purismo* y el *casticismo*. El purismo es una ideología que consiste en imponer a la cultura de la lengua modelos ejemplares, tomados de épocas de la historia de las lenguas que se lleguen a considerar de la mayor perfección y del mayor prestigio. En el caso del francés, el purismo se apoyaba en el aprecio de su literatura de los siglos XVII y XVIII; en el del español, en su literatura de los Siglos de Oro, entre el XVI y el XVII. Para la ideología purista, toda lengua, como los seres vivos, nace, se desarrolla, alcanza su mayor esplendor y después decae, se corrompe. El objetivo del purismo es lograr que los hablantes de una lengua se atengan únicamente al vocabulario y la expresión sintáctica cuyos modelos ejemplares se encuentran en sus épocas de esplendor; del español por ejemplo, en Cervantes, en Quevedo, en Lope de Vega, en Juan Ruiz de Alarcón, etc., con el objeto de detener su fatal corrupción. Todo vocablo nuevo que no pueda justificarse en el pasado dorado de la lengua es censurable.

Las ideologías puristas suelen acompañar épocas de decadencia, postración o confusión de las sociedades; por ejemplo, en España, el purismo tuvo su auge durante el siglo XVIII, como reacción al mal estado del reino y a la influencia del francés sobre su vida política y social. De esa época proceden obras como *Exequias de la lengua castellana* de Juan Pablo Forner. Puesto que, una vez enunciado el objetivo general, de sujetar la evolución de la lengua a los modelos ejemplares del pasado, es muy difícil que los hablantes se atengan a ellos, la acción del purismo es siempre defensiva y a posteriori. No puede ofrecer normas generales, sino sólo perseguir usos puntuales que ya se han dado. De ahí que el primer blanco de todas las censuras del purismo español hayan sido escritores del barroco, que se apartaban del ideal clásico, como Luis de Góngora, a quien acusaban de oscuro y difícil. La crítica purista se cernía sobre palabras utilizadas por aquellos escritores, como *ingurgitar*, *intersticios*, *ineluctable*, *intumescencia*, *funámbulo* o *torbellino* (salvo la última, todas ellas latinismos cultos); igualmente, sobre textos como el siguiente: "Dirige las rectitudes del juicio a las infalibilidades del ascenso" (una expresión que llenaría de orgullo a muchos políticos actuales).

Su segundo blanco fueron los *galicismos*, es decir, los préstamos del francés. *Tourbillón*, *resorte*, *remarcable*, *funcionario*, *tremar* por *temblar*, *en boga*, *bien entendido*, etc. son ejemplos de un conjunto bastante grande de galicismos que entraron a formar parte de la lengua española. Contra el galicismo, se defendía la necesidad de utilizar vocablos y expresiones españolas; es decir, depender solamente de la creatividad propia de las tradiciones verbales españolas. Conviene distinguir esta ideología endocéntrica, del purismo descrito antes, y llamarla *casticismo*. El "purismo" alemán de la época hitleriana era sobre todo un casticismo.

3.2. El casticismo

El casticismo es una fuerza activa surgida en España en la primera mitad del siglo XVIII, por acción del neoclasicismo, cuyo fin era defenderse de la influencia francesa, acudiendo a las fuentes de creación de la propia lengua española. Ideología defensiva también, se distingue del purismo por aceptar la necesidad de que se crearan vocablos nuevos, de acuerdo con las necesidades de la época, en particular, en el vocabulario culto y científico. Fray Benito Jerónimo Feijoo en el siglo XVIII y Miguel de Unamuno en el primer cuarto del XX fueron dos distinguidos impulsores del casticismo.

Purismo y casticismo siguen siendo ideologías defendidas por muchos individuos en las sociedades contemporáneas, aunque con objetivos diferentes. En México, a finales de 1982, la Comisión para la Defensa del Español (nótese la conjunción anglicista: debía ser "comisión de defensa del español") fue un caso característico de ideología purista, enderezada contra el inglés, que atacó también las tradiciones verbales populares y las jergas.

3.3. Las ideologías lingüísticas y las ideologías sociales

Es fácil comprender que las dos ideologías expuestas *basculan* (un galicismo) sobre valores ideológicos más amplios, como la defensa de la identidad social propia, el racismo, el nacionalismo, la moralidad definida por ciertos valores sociales e incluso la defensa de intereses de clase; hoy podemos agregar el etnicismo, el feminismo y la llamada "corrección política".

Ya vimos cómo el racismo, uno de los pilares del nacional-socialismo, acudió al casticismo lingüístico para tratar de extirpar de la lengua alemana toda influencia extranjera que atentara contra la "pureza" de la lengua y, con ella, contra la "pureza" racial. La moralidad sexual, por su parte, definió el vocabulario aceptable en sociedades como la victoriana de Inglaterra en el siglo XIX, dando lugar a considerar "impuras" muchas palabras, con lo que se abrió la puerta para que pulularan los eufemismos y las actitudes subversivas en los disfemismos. Cuando cierto grupo social logra imponer al resto de la sociedad sus valores, todos los vocablos que procedan de otros grupos serán objeto de censura, que es lo que sucede con las jergas o con los vocabularios rurales, a los que se tacha, por lo general, de "barbarismos y vicios". La actual "corrección política" es la que lleva a muchos gobernantes a fórmulas redundantes y poco económicas, como el tener que aclarar siempre "mexicanos y mexicanas", "profesores y profesoras"; el feminismo en Estados Unidos de América ha llegado a proponer sustituir la escritura de la palabra *woman* por "womin", alegando que, en la primera, hay tres letras —man— que contribuyen a sostener el predominio de los hombres sobre las mujeres; en español, el vocablo *género* se ha convertido en un sustituto de *sexo*, por influencia del inglés, en donde *género* no es término gramatical (pues el sustantivo inglés no marca el género) y porque *sex* se ha restringido a la acción sexual. Así, los estudios de la mujer se llaman en muchas universidades "estudios de género". En México, a los niños abandonados que viven en las calles se los describe como "niños en situación de calle".

El lingüista no deja de formar parte de una sociedad y vive sus dificultades, por lo que es difícil no verse atrapado por alguna de estas ideologías. Desde el punto de vista lingüístico, ninguno de los vocablos perseguidos es perjudicial para una lengua; en la mayor parte de los casos la enriquece. Los diferentes tipos de purismo son siempre empobrecedores; el casticismo, en cambio, puede ayudar a reactivar los recursos de formación de palabras en una lengua, siempre que no se ligue a otras ideologías, como el racismo o el etnicismo.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

tabú verbal; participación; motivación; arbitrariedad-convencionalidad; *thesei; physei;* eufemismo; disfemismo; norma social; norma lingüística; germanía; caló; argot; jerga; vocabulario de solidaridad; grosería; solecismo; barbarismo; purismo; casticismo.

OBRAS CONSULTADAS

CASAS, Miguel, *La interdicción lingüística. Mecanismos del eufemismo y el disfemismo*, Cádiz, Universidad de Cádiz, 1986.

LAMMERT, Eberhard, *et al.*, *Germanistik – eine deutsche Wissenschaft*, Francfort, Suhrkamp, 1967.

LARA, Luis Fernando, "El caló revisitado", en *Scripta philologica in honorem Juan M. Lope Blanch*, México, UNAM, 1992, t. II, pp. 567-592.

——, "La Comisión para la Defensa del Idioma Español de México: crónica de una política lingüística abortada", en *Políticas del lenguaje en América Latina*, número 29 de Iztapalapa, México, UAM, 1993, pp. 147-176.

LÁZARO CARRETER, Fernando, *Las ideas lingüísticas en España durante el siglo XVIII*, Madrid, CSIC, 1949.

PLATÓN, *Diálogos*, México, Porrúa.

11
LA ETIMOLOGÍA

INTRODUCCIÓN

Por su carácter de signo que refiere a objetos, acciones y relaciones, en el pensamiento de muchos seres humanos y de lingüistas que ignoran la teoría saussureana del signo la palabra se confunde fácilmente con aquello a lo que refiere. Ya vimos dos de las consecuencias de esa clase de concepción: por un lado, la confusión de la sustancia del contenido de la palabra con lo referido, o incluso con "conceptos" alojados en la mente, disuelve la significación en la referencia misma; para establecer la diferencia entre significado y referencia, de acuerdo con el pensamiento saussureano, nos extendimos en el tratamiento del significado de la palabra en el capítulo 4. Por otro lado, la idea de que el significado de la palabra participa de la esencia de los objetos significados, o tiene su motivación en la naturaleza de esos objetos, dio lugar, desde la Antigüedad, a la necesidad de investigar qué esencia o qué naturaleza de los objetos o las acciones referidos se esconde en el origen de las palabras.

En este capítulo exploraremos con cierto detenimiento la búsqueda de explicaciones del origen de las palabras, entendida en un principio como búsqueda de sus motivaciones en la naturaleza de las cosas, que es como comenzó la disciplina llamada *etimología*, pero que modernamente se ha convertido en una *historia de palabras*, en que confluyen la fonología y la morfología diacrónicas, la gramática histórica y todas las ciencias que pueden auxiliar al lingüista en el estudio de la aparición de las palabras en una lengua o en varias lenguas de la misma familia o del mismo tronco, como la historia, la arqueología, la escritura, la dialectología, la geografía lingüística y la historia literaria.

1. CONCEPTOS BÁSICOS

La etimología de *etimología* es el adjetivo ἔτυμος, que quiere decir 'verdadero o auténtico', y el sustantivo λόγος, que quiere decir 'palabra'. Según los antiguos griegos, la etimología es el estudio del origen verdadero o auténtico de una palabra. Vimos en el capítulo anterior que Sócrates proponía a Cratilo y Hermógenes que las palabras debían tener su motivación en una comprensión completa de la esencia de aquello que significan. Tal comprensión garantizaría que la palabra tuviera una relación verdadera con la cosa significada, lo que venía a ser un importante requisito para el discurso filosófico, en donde se espera que todas sus proposiciones sean verdaderas.

El diálogo *Cratilo* no nos ofrece, sin embargo, criterios o métodos para realizar esa investigación del origen auténtico de las palabras. Aunque verosímiles y, en muchos casos, encantadoras, las explicaciones etimológicas de Sócrates nos dejan con la duda de qué tan ciertas podrán ser. Veamos por ejemplo, la explicación de Sócrates de la etimología de *anthropos* 'ser humano': ἄνθρωπος debe provenir de la expresión ανα θρων α όπωπεν, que quiere decir 'contemplador de lo que ha visto' porque, a diferencia de los demás animales, que ven las cosas sin examinarlas ni contemplarlas —ἀναθρει—, cuando el hombre mira una cosa —ὀπωπε—, la contempla y se da razón de ella. ¿Cómo podemos saber si eso es cierto? ¿Cómo explicar la condensación de ανα θρων α όπωπεν en ἄνθρωπος, desde el punto de vista de su fonología y su morfología?

Una idea semejante a la de Sócrates en el *Cratilo* dio lugar al gran trabajo enciclopédico de San Isidoro, obispo visigodo de Sevilla, a principios del siglo VII. Su recopilación de cuanto pudo encontrar en el campo del saber de su época, llamada precisamente *Etymologiae* (y también *De origine*), suponía que el estudio del significado originario de las palabras llevaba a un conocimiento certero de las cosas. Como Sócrates, sólo que a gran escala y con una heterogénea documentación, pues las *Etimologías* constan de 20 libros, en que trata el *trivium* (gramática, retórica y dialéctica) y el *cuadrivium* (aritmética, geometría, música, astronomía) —base de la formación intelectual en la Edad Media— y una multitud de temas religiosos y profanos de su época, San Isidoro se esforzó por encontrar las causas de los significados de las palabras en la esencia de las cosas (etimología *ex causa*), y los orígenes o proveniencia de las palabras (etimología *ex origine*), convencido de que, cuando se conoce la etimología, se comprende mejor la realidad.

Comparemos, por ejemplo, la etimología de *homo* 'hombre (como ser humano)', que ofrece San Isidoro, con la de su correspondiente griego *anthropos*, que vimos antes: "Llamamos así al hombre (*homo*), porque está hecho de *humus* (barro), tal y como se dice en el *Génesis*". A falta de un estudio etimológico del griego y del latín, que confirmen esos orígenes, las etimologías isidorianas nos dejan más preguntas que respuestas. ¿Será posible, por ejemplo, como afirma San Isidoro, que la palabra *formōsus* —de la que viene el español *hermoso*— derive de *formum*, "que los antiguos empleaban para designar lo caliente", porque son hermosas las personas a quienes el calor de la sangre las enrojece?

Tanto en Sócrates como en San Isidoro podemos reconocer un principio lingüístico de interpretación etimológica, vagamente basado en la cercanía de las formas de la expresión de las palabras, aunque a veces tan fantasioso y caprichoso que se entiende la burla de Voltaire cuando decía que la etimología es una ciencia en la que las vocales no son nada y las consonantes bastante poca cosa.

Habría de consolidarse la lingüística comparativa en el siglo XIX y posteriormente la lingüística histórica para que la etimología se convirtiera en una disciplina lingüística, dotada, si no de un método rígido, sí de instrumentos conceptuales que permitieran construir hipótesis relativamente verificables acerca ya no de los orígenes de las palabras, y menos de su motivación en la naturaleza de las cosas, sino de la evolución de muchas de ellas de una lengua a otra, o de una época a otra. La lingüística estructural, al introducir las nociones de sistema y de campo léxico, es la que vino a dar a la etimología moderna sus características actuales.

La etimología creada por esta lingüística se concibe hoy en día como estudio de la *historia de las palabras*; podemos considerarla también como parte de una lexicología histórica, cuyo punto de partida consiste en proponer un vocablo como precedente, que constituye el *étimo* de otro, supuesto como derivado suyo, sobre la base del significante y el significado de ambos signos. Tal punto de partida supone que haya relaciones formales o culturales entre ambos vocablos, que permitan elaborar la hipótesis inicial, y también que se toman en cuenta todas las características de los sistemas lingüísticos puestos en relación; particularmente, que se conocen los procesos de cambio fonológico y morfológico entre uno y otro sistema.

1.1. El cambio fonológico

Entre un étimo E y la palabra considerada derivada suya D puede haber varios estadios evolutivos de una lengua, cada uno de ellos con características propias. El método etimológico da prioridad a la evolución fonológica; es decir, al cambio fonológico que se produce en una lengua, por ser este el fenómeno más regular y más sistemático; tal cambio es resultado de procesos que se repiten en la mayor parte de los signos en su evolución de E a D, como lo proponían las llamadas "leyes fonéticas" preconizadas por la escuela de los Jóvenes Gramáticos (*Jung Grammatiker*) o Neogramáticos (en la traducción que hizo del mote el lingüista italiano Ascoli), que a finales del siglo XIX se dieron a conocer desde Alemania.

Tal es el caso, por ejemplo, de la evolución del grupo /kt/ entre el latín y el castellano. Las consonantes evolucionan regularmente según sus características articulatorias, según si son simples o forman grupos consonánticos y según su posición en la palabra. En el paso del latín al castellano, los fonemas /k/ y /t/ que aparecen agrupados, pertenecen a sílabas distintas. En un principio el fonema /k/, final de sílaba, se vio atraído hacia el punto de articulación del fonema /t/, por lo que pasó de ser velar a pospalatal y luego se relajó para convertirse en fricativo prepalatal, dorsal y sordo [y°] en el llamado "latín vulgar" de la Alta Edad Media; este fenómeno, compartido por los demás dialectos peninsulares, se singularizó después en el castellano, porque el sonido prepalatal atrajo la articulación de /t/, para hacerlo dorsal y prepalatal, dándole un sonido "chicheante", que dio como resultado la formación de un solo fonema palatal africado sordo /ch/. Así, el latín *lactuca* dio *lechuga*, *factu* > *hecho* (nos concretamos al fonema /ch/), *lacte* > *leche*, *tectu* > *techo*, etc. La gramática histórica nos proporciona un análisis completo de esta clase de evoluciones.

1.2. El cambio morfológico

La consideración de la morfología del derivado D también suele conducir a una buena hipótesis etimológica, por cuanto la gramática histórica nos ofrece un panorama general de la evolución morfológica de las palabras en familias enteras de lenguas, como las romances o las germánicas. Es característica de las romances, por ejemplo, la procedencia de muchas palabras de formas diminutivas latinas, como *auricula* > *oreja*, *ovicula* > *oveja*, *clavicula* > *clavija*.

1.3. Los cambios originados en la vida social

Pero si el método etimológico da predominancia a las evoluciones sistemáticas, como condiciones *básicas*, muchas evoluciones dependen de factores sociales diversos, que requieren elaboraciones mucho más detalladas y bien documentadas. Tomemos por caso la palabra *hermoso*, que San Isidoro trata tan simpáticamente. El paso de *formōsus* a *hermoso* en español es regular: el fonema /f/ inicial se aspira en castellano por influencia del vasco, que no tiene un fonema semejante; los dos fonemas /o/ pretónico y tónico se disimilan y la terminación en /u/ pasa a /o/ como parte de la simplificación de las vocales finales del latín al español, que producen solamente /a, e, o/. En portugués y rumano *formoso* y *frumos* proceden de la misma palabra latina.

En cambio, en francés y en italiano, las palabras correspondientes provienen de *bellus* 'bello': *beau/belle* y *bello* (que en español tengamos también *bello* se debe a una adaptación del antiguo occitano, que se introdujo al castellano más tarde, en el siglo XIII) ¿Qué explicación se puede encontrar para ello? Una tesis general sobre la evolución de las lenguas romances, elaborada por los lingüistas italianos Matteo Bartoli y Giulio Bertoni, basada en la corriente llamada "lingüística espacial" o "neolingüística", de la que ellos fueron impulsores, propone que la evolución del latín en sus territorios europeos permite distinguir una zona central, formada por Italia, sobre todo en su parte norte, y la Galia (quizá incluso los territorios catalanes), y una periferia, formada por la península ibérica y Rumania en sus extremos, pero también el sur de Italia y Sicilia. *Bellus* correspondería a una etapa más moderna del latín central, en tanto que *formōsus* sería una forma más antigua, conservada en su periferia.

El gran filólogo Yakov Malkiel relata que hay pruebas documentales de que *formōsus* es anterior cronológicamente a *bellus*, por lo cual la hipótesis de la lingüística espacial parece confirmarse. Por cierto, *formōsus* proviene de *forma* —lo que también señalaba San Isidoro, antes de

inclinarse por su hipótesis "térmica"— en tanto que la segunda viene de *bonus* 'bueno'.

No es raro que la historia de las palabras muestre irregularidades que no se producen en la evolución de un sistema lingüístico, sino por diferentes acontecimientos en las sociedades. Alberto Zamboni, autor de una valiosa exposición de nuestro tema, nos ofrece el siguiente ejemplo: el vocablo latino *cŏxa* 'cadera' no tuvo sucesión en italiano, francés, español y portugués, en donde se vio sustituido por *anca* (it.), *hanche* (fr.), *anca* —para los caballos— y *cadera* —para las personas, tomada a su vez del latín *cathĕdra* 'silla'— (esp.), y *anca* (port.) respectivamente. *Anca* es una adaptación de una palabra germánica **hanka*. En cambio, el significante *cŏxa* evolucionó a *coapsă* en rumano, *cueisa* en provenzal y *cuxa* en catalán, pero con el significado de 'muslo'; en italiano también *coscia* pasó a significar 'muslo'. Walter von Wartburg, autor de un monumental diccionario etimológico del francés (*Französisch Etymologisches Wörterbuch*), sostiene que todo eso se debe a una homofonía que se había producido en latín posclásico, entre *fēmur* 'muslo' y *fĭmus* 'estiércol'.

Por ese carácter simbólico de las palabras que tratamos en el capítulo anterior, ante el peligro de confundir el muslo con el estiércol llevó a los hablantes de esas regiones a "correr" el significado de *cŏxa* a 'muslo', para eliminar *fēmur*, y sustituirlo, en español, por *muslo* (del latín *mūscŭlum* 'ratoncito'), en italiano, como ya vimos, por *coscia*; en francés por *cuisse* y en portugués por *coxa*. (*Fémur* volvió a estas lenguas como cultismo, introducido por la anatomía.) Al quedar la cadera sin nombre, se abrió la puerta para la adopción del germánico **hanka* como su nombre (y, en español, el cambio de *cathĕdra* 'silla' a *cadera*).

El asterisco que precede a una forma lingüística —como en **hanka*— es una convención gráfica introducida en lingüística por August Schleicher, en su *Compendium der vergleichenden Grammatik der indogermanischen Sprachen* (1861-1862), para señalar que se trata de una hipótesis, sin sustento documental.

1.4. Préstamos de otras lenguas

Si en el ejemplo anterior vimos cómo un préstamo germánico llegó a ocupar un lugar entre los nombres del cuerpo humano, este fenómeno es muy usual debido al contacto que se produce entre las lenguas. Tomemos por caso los préstamos del árabe, que fue, durante casi 800 años,

una lengua de civilización tan importante como el latín (y en varios campos, más que el latín) en la península ibérica. Palabras como *almohada* (<*muhadda* < *hadd* 'mejilla'), *albañil* (< *banna'* 'constructor'), *alberca* (< *birka* 'estanque'), *álgebra* (< *gabr* 'reducción') tienen su etimología en la lengua árabe (en todas ellas, el español conserva el artículo árabe).

Sin embargo, atribuir a un contacto de lenguas y a un préstamo el origen de una palabra, como primer paso de la investigación etimológica, puede dar lugar a errores. Por ejemplo, *almena*, aunque a primera vista parezca formar parte de la herencia árabe, es de origen latino: *mĭna* 'saliente' (después se le añadió el artículo árabe). Lo mismo sucede con *almendra*, del latín vulgar *amyndula* y ésta del clásico *amygdala*, tomada del griego. Por ese motivo, el método etimológico prefiere dejar en lugar secundario la posibilidad de un préstamo de otras lenguas en el origen de una palabra, hasta después de haber agotado los criterios sistemáticos.

Esta clase de prudencia es especialmente necesaria cuando, muchas veces por motivos patrióticos o incluso por rivalidades entre etimólogos, se tiende a privilegiar las hipótesis de préstamo sobre todas las demás. Eso sucede, por ejemplo, en México, en donde es motivo de orgullo la etimología prehispánica de muchas palabras: *tomate*, *aguacate*, *chocolate*, *cacahuate*, *chilpayate*, *escuincle*, etcétera.

Veamos tres ejemplos de esta clase de equivocaciones: en 1943, Alfredo Barrera Vázquez, uno de los más famosos mayistas mexicanos, propuso que el vocablo *apesgar*, que en el español yucateco quiere decir 'ejercer presión sobre algo, poner un peso encima de algo' es un préstamo del maya *petsah* 'presionar, apretar'. Sin duda hay alguna similitud fonética entre ambos vocablos.

El método etimológico exigiría: *1*] revisar antes todos los casos probados de etimología maya de préstamos al español, para conocer 2] qué procesos de adaptación fonológica se produjeron y 3] qué características morfológicas y sintácticas tiene la palabra en la lengua de partida, para poder plantear una hipótesis sobre la manera en que se insertó en español, pero además exige 4] revisar toda la bibliografía previa que sea posible, para cerciorarse de que el vocablo no tiene documentación *a*] fuera de la zona de contacto maya, *b*] en el español nacional de México, *c*] en otros dialectos del español. Yakov Malkiel, uno de los más grandes etimólogos de las lenguas romances, demostró que *apesgar* está docu-

mentado en el occidente de León, España, y que *pesga* 'peso' tiene una documentación tan antigua como el siglo XIII, procedente del latín *pēnsum* 'peso'. *Apesgar*, en consecuencia, no es un mayismo del español de Yucatán.

Tal error no invalida todo el trabajo de un etimólogo, como lo demuestra el mismo Malkiel, al corroborar que el uso yucateco de *acechar* 'visitar a alguien de improviso, "caerle" a alguien (como dicen los jóvenes mexicanos)' no tiene documentación en otros dialectos del español y discrepa notablemente del significado principal del vocablo 'esperar la oportunidad para atacar a alguien'. Así, sólo en español de Yucatán se dice: "fui a *acechar* a mi tía cuando estuve en el pueblo". El maya *ch'eneb* tiene ese significado. Barrera Vázquez estaba en lo correcto. Hay que notar que el préstamo es sólo de sustancia del contenido, no de forma.

Acostumbrados a que vocablos como *chocolate*, *tomate*, *zapote*, *cacahuate*, etc. tienen una evolución regular en su paso al español como préstamos nahuas: *xocolatl* > *chocolate*, *tomatl* > *tomate*, *zapotl* > *zapote*, *cacahuatl* > *cacahuate*, es muy sencillo suponer que todas las formas terminadas en *-ate* sean nahuatlismos. Juan M. Lope Blanch nos demostró que *ate*, el nombre de esa rica pasta dulce de frutas, como el membrillo o la guayaba, que se fabrica sobre todo en Michoacán, no forma parte de aquel sufijo: hay registro de *piñonate* en España desde 1525, de *codoñate*, mucho antes (en aragonés, desde 1373) y además *avellanate*, *almendrate*, *membrillate* e incluso como nombres de guisados: *higate*, *romerate*. El origen de ese sufijo *-ate* para los dulces es catalán: *codonyat*, *pinyonat*, castellanizado posteriormente. En cambio, su conversión en palabra *-ate-* es mexicana.

Joan Corominas, a quien le debemos el único gran diccionario etimológico del español, a su vez, también se dejaba llevar a veces por su españolismo o por tratar con desdén el supuesto origen amerindio de muchos vocablos. Tal es el caso de la etimología de *tocayo*, con que en México y en otros dialectos del español llamamos a quien tiene nuestro mismo nombre de pila. En náhuatl, *tocaitl* quiere decir 'nombre' y *tocayotia* 'llamar a alguien por su nombre'. A pesar de eso, Corominas prefirió dar como explicación de su etimología una fórmula ritual del casamiento romano, según la cual, cuando la novia llegaba a la casa del novio decía: "*Ubi tu Cajus, ibi ego Caja*" (en donde tú estás, Cayo, estoy yo, Caya). La fórmula no se utilizó más entre los pueblos de lengua romance, en donde el cristianismo había desterrado la ritualidad pagana; pero además, es difícil que "tú, Cayo", se sustantive en *tocayo*.

El desdén de Corominas por lo americano no daña gravemente la importancia y la calidad de su diccionario, que es el único actual que se puede recomendar. En varios casos, más que desdén parece haber privado en él la falta de datos. Por ejemplo, hay dos homónimos *biznaga* en español: uno, que nombra a una planta umbelífera, de la misma familia que la zanahoria, cuyo nombre proviene de *bišnâqa*, una adaptación que hizo el mozárabe del latín *pastinaca*; otro, adaptado por el español del náhuatl *huitznáhuac*, que quiere decir 'rodeado de espinas', y nombra a varias especies de cactáceas, de forma globosa y piel llena de espinas, de una de las cuales se obtiene el *acitrón* (que a su vez, es homónimo de otro, español, nombre de un cítrico). Quizá le faltó conocer esta información.

Se vuelve más complicada la demostración etimológica cuando hay explicaciones sistemáticas para la proveniencia de una palabra, e incluso hay documentaciones que parecen confirmarlas, pero el funcionamiento y el significado de la palabra en cuestión no corresponden al sistema que la podría explicar. Ese es el caso del morfema *-le* en español mexicano, en expresiones como *ándale*, *córrele*, *apúrale*, *vuélale*, todas ellas exhortativas. El morfema de complemento indirecto *-le* del español, en expresiones como *dile*, *entrégale*, *mándale*, e incluso en las expresiones leístas del español peninsular *salúdale* (en vez de *salúdalo*), *llámale* (en vez de *llámalo*), *bésale* (en vez de *bésalo*), *invítale* (en vez de *invítalo*), etc. han parecido a muchos prueba suficiente de que *ándale*, *córrele*, etc. del español mexicano son manifestaciones de ese morfema. Sin embargo, en el uso mexicano *-le* se adhiere a formas imperativas intransitivas, es decir, que no rigen complemento indirecto. En cambio, sabemos que en náhuatl hubo una interjección exhortativa *áxcalli*, que dio en español *áxcale* (ya en desuso), a la que la similitud con el morfema del español terminó por segmentar: *-le*, que resulta homónimo del anterior.

1.5. El fonosimbolismo

En el *Cratilo*, Sócrates propone que en el origen de los "nombres primitivos", o sea, de las palabras más antiguas de una lengua (en este caso, del griego) no está "la imitación de un objeto con la voz", es decir, no es-

tá una *onomatopeya*, sino la imitación mediante los sonidos de la esencia de las cosas. Así, el fonema /r/ del griego se presta para significar el movimiento: ῥεῖν 'correr', ῥοή 'curso', τρόμος 'temblor' (lo mismo diríamos del fonema /r/ en español; recordemos el ejercicio para ejercitar la pronunciación de /r/ a los extranjeros: "rápido ruedan las ruedas del ferrocarril"); el fonema /i/, dice Sócrates, sirve para significar lo sutil: ἱέσθαι y ἴέναι, las acciones de ir y marchar, etc. A esta posibilidad se le llama hoy en día *fonosimbolismo*.

Incluso la terminología fonética se deja llevar por el fonosimbolismo: a las vocales se les llama "agudas" u "oscuras"; hay consonantes "chicheantes", "ásperas", "blandas", etc. A lo pequeño, por ejemplo, muchas lenguas le dan signos en que predominan las íes, como indica Otto Jespersen, el lingüista danés que tanta influencia tuvo, sobre todo, en Estados Unidos de América: en inglés, *little*, *tiny*, *weeny*, *slim*, *bit*, *tip*, *slip*; en danés, *bitte*; en latín, *minor*, *minimus*; en italiano, *piccolo*, *piccino*; en español, *chico*, *fino*. Con íes se construyen muchos sufijos diminutivos: *-y*, *-ie* en inglés; *-i* en húngaro; *-ino* en italiano, *-ito*, *-ico*, *-illo* en español; *-tzin* (diminutivo y honorífico) en náhuatl, etc. Al valor simbólico del diminutivo podemos agregar hipocorísticos como *Güicho* (Luis), *Fito* (Rodolfo o Adolfo), *Guille* (Guillermo) o, en inglés, *Eddie*, *Tommy*, *Billy*, *Jeannie*, *Cathy*, etcétera.

El fonema /ch/ en español es especialmente preferido para formar hipocorísticos, palabras festivas e interjecciones: *Chela* (Marcela), *Chabela* (Isabel), *Charo* (Rosario), *Chelo* (Consuelo), *Chente* (Vicente), *Chole* (Soledad), *Pancho* (Francisco), y *chiquis*, *chato*, *chino*, *chacho*, según ciertos rasgos físicos y de edad. ¡*Chin*! (apócope de *chingados*, pero de valor eufemístico) y ¡*chale*! exclamativos; ¡*Ya chole*! para deshacerse de una molestia, basan su expresividad en ese fonema.

1.6. La onomatopeya

Tema predilecto de la fantasía etimológica y de quienes creen que en el origen de las lenguas hay una imitación de sonidos de la naturaleza, la onomatopeya tiene un papel en la formación de muchas palabras, pero no puede considerarse el origen principal de los léxicos de las lenguas humanas (al menos, no según los conocimientos actuales).

Entendemos por *onomatopeya* toda palabra que, en su expresión, imita un sonido de la naturaleza. A diferencia del fonosimbolismo, que tiene una base *sinestésica*, es decir, de interpretación de un estímulo complejo, como la percepción de lo pequeño o la percepción del movimiento, mediante el sentido del oído, la onomatopeya consiste en una adaptación del sonido escuchado a las características del sistema lingüístico de cada lengua histórica. Así por ejemplo, el canto de los gallos se traduce en signo, en español, como *quiquiri-quí*, mientras que en francés es *cocorico* y en inglés *cock a doodle do*; los gatos hacen *miau* en español, pero *meouw* en inglés; los perros hacen *guau* en español, pero *wuf* en alemán.

En cuanto las palabras de origen onomatopéyico son signos de las lenguas, sufren la evolución del resto del léxico y de esa manera pierden la "transparencia" del sonido original. Por ejemplo, *vagir* en español, no deja ya translucir su origen en una onomatopeya latina *vagīre*, pronunciada [wa'gi:re], que imitaba el [wa] del borrego (en español, los borregos hacen *beé*). Una raíz indoeuropea **kleg-*, considerada onomatopéyica del ruido que produce la risa, dio en antiguo alto alemán *hlan hlan*, que luego evolucionó a *lachen* 'reír', en el moderno alto alemán; *helahhan hliehhan* en anglosajón, que dio *laugh* en inglés.

El inglés es especialmente propenso a formaciones onomatopéyicas, como *bash* 'golpe', *smash* 'golpe seco', *crash* 'choque', *lash* 'latigazo', *splash* 'chapuzón' o *flash* 'relámpago' (de ahí una de las dificultades de la traducción técnica del inglés al español, pues muchos términos técnicos ingleses son onomatopéyicos o de base fonosimbólica, como *flip flop* 'vuelta de 180° o hasta chancleteo', *beep* 'silbido o pitido', *boom* 'auge', *pop* 'reventar', lo que muchas veces imposibilita encontrar una palabra en español que sea tan plástica y tan corta). En cambio, si consideramos las llamadas "voces" de los animales, en *barritar* (de los elefantes), *relinchar* (de los caballos), *rebuznar* (de los burros), *arrullar* o *zurear* (de las palomas), *garrir* (de los loros), no podemos reconocer en ellas alguna base fonosimbólica u onomatopéyica, como sí lo hacemos en *piar*, *gruñir*, *mugir* o *berrear*.

2. LA ETIMOLOGÍA POPULAR

Esa humana y persistente búsqueda de motivación para las palabras, que manifiestan el diálogo *Cratilo* y las *Etimologías* de San Isidoro da lugar a permanentes intentos por encontrar sus etimologías. Basta un pa-

recido fonético entre dos signos para que la fantasía se apreste a encontrarles una relación. Cuando se produce este fenómeno, el esfuerzo de las personas da lugar a lo que llamamos *etimología popular*.

Veamos algunos casos repetidos de etimología popular en México: mucha gente sostiene que *gringo*, el gentilicio despectivo que utilizamos los mexicanos para hablar de los naturales de Estados Unidos de América (y los argentinos para hablar, en especial, de los italianos) nació durante la invasión estadounidense a México en 1847, porque los soldados animaban a sus caballos a comer el pasto verde diciéndoles: ¡*green, go*! (como si un caballo necesitara que lo animaran a comer). Corominas nos demuestra que *gringo* ya se usaba en el sur de España desde el siglo XVIII para denominar a los extranjeros, como modificación fonológica de *griego* (quizá una velarización del diptongo, que luego se oyó como nasalización) pues "hablarle a uno en griego" quería decir "hablarle en una lengua incomprensible" (hoy decimos "hablarle a uno en chino"). La etimología de *gringo* es, pues, *griego*. Por eso en Argentina se llama así a los extranjeros y, en México, se ha especializado para los estadounidenses.

Otra repetida etimología popular es la de *mariachi*, que se ha querido retrotraer al francés *mariage* 'casamiento', suponiendo una influencia francesa de la época de Maximiliano, en el español del occidente de México. Pero hay documentación de la palabra desde 1852 (antes, por lo tanto, de la intervención francesa en México) que muestra que *mariachi* era, primero, la tarima en que se bailaban el son o el fandango; después, los músicos que tocaban esos sones jaliscienses, e incluso, a veces, los sones mismos. Podemos estar seguros, en consecuencia, de que la etimología propuesta **mariage* > *mariachi* es falsa, aunque hasta la fecha no se haya podido proponer una hipótesis seria de la etimología de *mariachi* (se duda de las propuestas de que está tomada de una "lengua coca" del occidente de México, o del pinutl).

La expresión técnica *etimología popular* fue introducida por primera vez en lingüística por el germanista alemán Ernst Förstermann, en 1852. De esa manera trataba de distinguir la "etimología culta" (*gelehrte Etymologie*; hoy diríamos "etimología científica") de la que producen los intentos populares por motivar varias de sus palabras.

Por último veamos el caso de la palabra de tradición popular mexicana *abusado*, que quiere decir 'listo, alerta, astuto' y que se ha querido explicar como modificación de *aguzado* 'que tiene punta, que es agudo'. Su etimología es *buzo*, un vocablo de la germanía española que quiere

decir, precisamente, lo mismo que *abusado*. Cuando en México se usa la variante "¡ponte buzo"! en vez de *abusado*, se está reactivando la primitiva voz germanesca.

La etimología popular sería inexplicable, como fenómeno lingüístico, sin la necesidad humana de encontrar motivación en sus palabras. Aparte de que suele ofrecer explicaciones fantasiosas muy divertidas, para la lingüística es un fenómeno de dos caras: por un lado, permite estudiar la manera en que los seres humanos se esfuerzan por encontrarle alguna justificación a sus palabras; por el otro, es la clase de fenómenos que daña la investigación científica, pues le abre la puerta a muchos aficionados que, sin preparación ninguna, acometen la búsqueda de etimologías y confunden a la sociedad.

3. LA ONOMÁSTICA

Parte de los estudios etimológicos corresponde a una disciplina desairada por la lingüística moderna que, sin embargo, constituye uno de los temas de interés de la vida social. Tal disciplina es la *onomástica*, o estudio de los nombres propios, de los nombres de habitantes de países, regiones, pueblos e incluso barrios de ciudades (*antropónimos*), de los nombres de países, regiones, pueblos y lugares (*topónimos*).

La búsqueda de motivos o razones para nombrar de cierta manera a una persona o a un lugar no sólo interesa a los padres de familia que desean dar a sus hijos nombres simbólicos de personajes apreciados, como Orestes, Aquiles, Virgilio, Julio César, Dante, Alfonso, Rodrigo, Erasmo, Juan Luis, Juan Sebastián, Martín Lutero, Fidel, Lenin, Cuauhtémoc, etc. (una tarea que acometen los almanaques o los libros como el *Diccionario etimológico comparado de nombres propios de persona*, de Gutierre Tibón, 1986), sino a la investigación histórica, que obtiene de los nombres de personajes y de lugar datos importantes para la historia de los pueblos.

Sócrates afirmaba en el *Cratilo* que Afrodita la diosa griega de la belleza, Venus para los romanos (véase el cuadro de Botticelli), tenía ese nombre porque nació de la espuma del mar, es decir, τοῦ ἀφροῦ, lo cual confirmaba el mito griego de su nacimiento; hoy se sabe que el nombre es el de una divinidad del Medio Oriente, de origen fenicio: *Aštoreo*,

Aphroteo. Esta clase de justificaciones a posteriori de los nombres de personas o de divinidades ofrece un dato importante para la historia de las religiones y de las culturas.

Por la expresión de los nombres se forman justificaciones como Santa Lucía como protectora de los ojos (de *luz*) en español e italiano, una tarea de Sankt Augustin entre los alemanes (por *auge* 'ojo'); en francés, Saint René es el protector de los riñones (*reins*). El valor simbólico de las palabras hace que un lugar como *Maleventum*, una ciudad italiana de origen griego, se cambie por *Beneventum*.

El caso de *Cuernavaca*, en México, señala la manera en que los conquistadores españoles interpretaban las palabras nahuas (*Cuauhnáhuac* 'rodeada de árboles'). El estudio toponímico del vocablo catalán *Ultramort* (¡más allá de la muerte?) lleva a reconocer un originario *Ultramont* 'más allá de los montes'; en las faldas del volcán Etna, un pueblo se llama *Linguaglossa* (*lingua* en latín y *glossa* en griego significan 'lengua'), una modificación de su nombre originario *Lingua grossa* 'lengua grande', que nombra una gruesa lengua de lava que emanó del volcán, y documenta un fenómeno fonético llamado *labdacismo* o mutación de [r] a [l]. El nombre de la constelación que llamamos Osa Mayor proviene de un error de interpretación del asirio *ereqqu* 'carro', cuya raíz *r-q-q* (recordemos que en las lenguas semíticas las vocales varían según la flexión) llegó a sonar a oídos griegos *árktos* 'oso'.

La *toponimia*, o estudio de los nombres de lugares, que forma parte de la onomástica, tiene una importancia central para la geografía, para la arqueología, para la historia y para la historia de las lenguas, pues en muchos casos los ríos, las montañas o los pueblos conservan nombres que corresponden a épocas históricas muy antiguas, que sirven como indicios de la existencia de diferentes sociedades y de diferentes lenguas. En España, por ejemplo, son los topónimos los que ayudan a sostener que el norte de la península conserva mejor los indicios de épocas más arcaicas, que el sur, durante 800 años territorio musulmán.

Por los estudios toponímicos se afirma que no hay un sustrato ibérico único en la península, sino que convivieron varios pueblos diferentes antes de la colonización romana. Indican presencia vasca los nombres de *Lascuarre* y *Loarre* en Huesca (Aragón); *Gerri* y *Algerri* en Lérida (Cataluña); céltica en *Segovia* < *Segobriga*, *Coimbra* < *Conimbriga*, y en la cuenca del río Ebro, como *Berdún* y *Asún* en Huesca, *Navardún* en Zaragoza; son de origen fenicio *Cádiz* < *Gadir* 'fortaleza', *Málaga* < *Malaca* 'reina'; es griego *Ampurias* < *Emporion* (*Empuries*, en catalán) ; son de origen árabe *Gibraltar* < *Gibr al Tarik* 'peña de Tarik', *Guadalajara* < *Guad al aixara* 'río de las piedras'.

Tratándose de México y, en general, de Hispanoamérica, en donde hubo muchos cientos de lenguas aborígenes antes de la Conquista y colonización española, y en donde todavía hay varios cientos de diferentes lenguas, la toponimia es un campo abierto a la investigación, que no sólo interesa al conocimiento, sino también a las agencias y a los ministerios de gobierno que tienen que producir buenos mapas, sustento geográfico fidedigno al levantamiento de censos, dirimir conflictos de límites entre municipios y localidades, etcétera.

La toponimia prehispánica enfrenta la dificultad de la inexistencia de registros más antiguos que el siglo XVI, que ayuden a construir hipótesis acerca de la etimología de cada topónimo. Los jeroglíficos que se han ido identificando en varias regiones de Mesoamérica, que simbolizan significados de pueblos mayas, nahuas, zapotecas, mixtecas, etc. no permiten interpretaciones unívocas debido a su carácter ideográfico de base, aunque, como vimos al comienzo de este *Curso*, contengan también símbolos de interpretación silábica. Eso dificulta mucho la tarea de los etimólogos dedicados al estudio de la toponimia mexicana, pero a la vez es una invitación a emprender esa clase de aventura intelectual.

Algunos topónimos son de fácil comprensión, como *Chapultepec*, o 'cerro del chapulín (cerro del grillo)', cuyo jeroglífico náhuatl es muy plástico. Acerca de *Tenochtitlan*, nombre primitivo de la Ciudad de México, ha habido un largo debate acerca de su interpretación, pues si *tenochtli* quiere decir 'tuna (higo chumbo)' —pero se toma como metáfora del corazón— y *tetl* quiere decir 'piedra', *Tenochtitlan* querría decir 'piedra de la tuna'; el jeroglífico, que dio origen al escudo nacional, tiene sobre todo un carácter simbólico mítico, por lo que hay quien interpreta todo el conjunto como símbolo de un sacrificio ritual, realizado sobre la piedra que emerge de la laguna. El nombre de *México* ha dado lugar a varios tratados; la etimología propuesta es 'en el ombligo de la Luna', formada por *meztli* 'luna', *xictli* 'ombligo' y *co* 'lugar'.

La etimología en lenguas amerindias todavía presenta grandes dificultades de interpretación, debido a la falta de documentos que ofrezcan datos acerca de la probable evolución de estas lenguas en el milenio

anterior al siglo XVI. A ello hay que agregar las características propias de cada lengua, como en el caso del náhuatl, que es aglutinante por lo que, a la vez que facilita varias segmentaciones morfológicas posibles, el significado metafórico de muchas de sus palabras, relacionado de manera compleja con jeroglíficos, no ayuda a decidir cuál de ellas es la más verosímil.

4. CONCLUSIÓN

La etimología es, entre las disciplinas de la lingüística, la que más requiere tratamiento filológico; es decir, en la que hay que poner en juego no sólo el instrumental sistemático de la lingüística, sino el conocimiento de la historia, la literatura, la arqueología, la comprensión de los sistemas de escritura y cuanto dato pueda ofrecer la memoria humana. Ya hemos visto que, aunque se puede enunciar cierta jerarquía entre los conceptos del método etimológico, no se puede convertir en una mecánica, sino que hace falta prudencia e, incluso, templanza. Base de una lexicología histórica, relaciona de inmediato el rigor de la lingüística con la comprensión de los fenómenos históricos, a la vez que completa la unión entre la lengua y la cultura.

Al concluir la lectura de este capítulo, el lector deberá dominar los siguientes conceptos básicos:

etimología; étimo; evolución fonológica; evolución morfológica; gramática histórica; fonosimbolismo; onomatopeya; etimología popular; onomástica; antropónimo; topónimo; toponimia.

OBRAS CONSULTADAS

ALATORRE, Antonio, "Sobre americanismos en general y mexicanismos en especial", *Nueva Revista de Filología Hispánica*, 49,1 (2001), pp. 1-49.

COROMINAS, Joan, y José Antonio PASCUAL, *Diccionario crítico etimológico castellano e hispánico*, Madrid, Gredos, 1980.

GUZMÁN, Ignacio (ed.), *De toponimia y topónimos. Contribuciones al estudio de nombres de lugar*, México, Instituto Nacional de Antropología e Historia, 1987.

LOPE BLANCH, Juan M., "Un falso nahuatlismo", *Nueva Revista de Filología Hispánica*, 27,2 (1978), pp. 296-298.

MALKIEL, Yakov, "On analyzing hispano-maya blends", *International Journal of American Linguistics*, 14,2 (1948), pp. 74-76.

——, *Etimología*, Madrid, Cátedra, 1996.

MENÉNDEZ PIDAL, Ramón, *Manual de gramática histórica española*, Madrid, Espasa Calpe, 1962.

ROHLFS, Gerhard, *Antroponimia e toponomastica nelle lingue neolatine. Aspetti e problemi*, Tubinga, G. Narr, 1985.

SEVILLA, Isidoro de, *Etimologías*, edición de José Oroz R. y Manuel-A. Marcos Casquero, Madrid, Biblioteca de Autores Cristianos, 2000.

ZAMBONI, Alberto, *La etimología*, Madrid, Gredos, 1988.

ÍNDICE ANALÍTICO

acción, acto verbal y: 41-43, 106; y el acto verbal de denominación: 43-48; esquemas de: 92, 100
acento, circunflejo: 122, 125-126; diacrítico: 122; función distintiva del: 59-60; de intensidad: 33-34; ortográfico: 30, 120-121, 125; prosódico: 30, 120
acentuación, reglas de: 121
actividad, y producto: 39-41
acto verbal, acción y: 41-43; condiciones del: 42; de denominación: 43-48; inteligibilidad del: 43
adaptación: 223; *véase también* extranjerismos
adjetivo: 38, 128-129
afasia, lenguaje infantil y la: 44
afijos, derivativos: 17; de flexión: 17
Alarcos Llorach, Emilio: 31-32
alemán, fonema fricativo: 32-33; germanización del: 225, 227; morfemas: 76-77; núcleos morfemáticos: 70, 73; onomástica: 244; onomatopeya: 241; unidades de denominación: 47, 72; verbos: 94-95
alfabeto: 23-24; fonético: 157
alma, potencias del: 54
ambigüedad: 106, 108; confusión entre vaguedad y: 107
análisis, binario: 192, 199; en conglomerados: 176-178, 183; cuarto nivel de: 80, 82; estructural de los signos: 181; etimologización del: 76; fonológico: 26-34, 184; preparación del: 185; primer nivel de: 21, 25; procedimiento de: 193-195; quinto nivel de: 82, 85; segundo nivel de: 23-34; semántico del léxico: 181, 185, 188, 191; semasiológicos: 205; tercer nivel de: 48
antonimia: 106; definición: 107
antonímica: 98
antónimos: 203
antropónimos: 107, 243
aprender, como aprender a hacer: 12
árabe: 116, 222; préstamos al español: 236-237
arbitrariedad-convencionalidad: 216
argots: 220-221
Aristóteles: 216; *Sobre la interpretación*: 38
artículos: 79, 137
asociación libre: 181, 183
Atlas lingüístico etnológico de Andalucía: 207
Atlas lingüístico etnológico de Colombia: 207
Atlas lingüístico de México: 207, 214
Austin, John L.: 42
Ávila, Raúl: 119

Ball, Hugo: 24
barbarismos: 224-225, 228
Barrera Vázquez, Alfredo: 237-238
Bartoli, Matteo: 235
base de la comparación: 60, 182
Bertoni, Giulio: 235
binario, análisis: 192-193
binarismo: 193
Bloomfield, Leonard: 24
Boas, Franz: 39
Bréal, Michel: 205
búlgaro, lengua: 117
Busa, Roberto, *Index thomisticus*: 163
bustrofedón: 119

calcos: 223; *véase también* extranjerismos
cálculo onomasiológico: 208-210
caló: 220-221, 225
campo asociativo: 182, 184; del, al campo léxico: 192-199; formación del: 183
campo léxico: 233; definición: 197; reestructuraciones del: 197-198
campo semántico, *véase* campo léxico
capacidad de sustitución paradigmática: 70
castellano, comparado con el latín: 38
casticismo: 226-229
Catach, Nina: 121
catacresis: 108
Cela, Camilo José, *Diccionario secreto*: 225
centro de interés: 173-175, 177
cirílico, escritura: 117
clasemas: 198-199
códices: 127
cohesión entre morfemas: 67
color focal: 90
Comisión para la Defensa del Español: 227
comparación, base de la: 60, 182, 195, 205
competencia léxica, disponibilidad y: 175-176
complexión, reglas de: 115, 120-127
composición de palabras: 74
comunidad de comunicación: 187
concordancias: 163
configuraciones: 88; morfológicas: 121-123; de carácter perceptual: 92, 95, 105
conglomerados, análisis en: 176-178, 183
conjunciones: 79, 137
conmutación: 59, 62, 68, 75; método de la: 24-27, 181-182, 184, 199; entre morfemas: 77; entre signos: 40
conocimiento, esquemas de: 92, 99; esquemas de, senso-motor: 44, 86-88; formas de: 44
conservación etimológica: 126
consonantes, alveolares: 27-28; ásperas: 240; blandas: 240; chicheantes: 240; dentales: 27-28; glotalizadas: 158; labiales: 27-28; oclusivas: 30, 32; palatales: 27-28; velares: 27
contenido, forma del: 57; expresión y: 53-57; sustancia del: 57
convención lexicográfica: 139
Corominas, Juan: 238-239, 242
corpus de datos, cualitativos: 161-162; de estudio del léxico de grupos sociales: 159-160; de finalidad exhaustiva: 162-163; formación de un: 153-154, 166; de ocurrencias: 169; suficiencia del: 157
Corpus de referencia del español actual (CREA): 162
Corpus del español mexicano contemporáneo: 156, 162, 169-170

Corpus diacrónico del español (Corde): 162
Correas, Gonzalo: 119
Coseriu, Eugenio: 182, 192
Covarrubias, Sebastián, *Tesoro de la lengua castellana o española*: 215
Cratilo: 232
cuadrivium: 232
cuadro fonológico: 27-28
cuestionarios: 161
cultura, definición de: 100; y la formación del significado: 99-103, 111; verbal: 101-102, 104-105, 139-140

Champollion, Jean-François: 26
Chang Rodríguez, Emilio: 169; *A Frequency Dictionary of Spanish Words*: 168
checo, acento de intensidad: 34
chino, escritura: 113
chol: 27

danés, fonosimbolismo: 240
Dante Alighieri, *De vulgari eloquentia*: 38
declinaciones: 128; en latín: 139
deixis: 208; local o mostrativa: 209; personal: 209
denominación, acto de: 43-48; unidades de: 47-48, 61-63, 66, 69-70, 72-75, 78-79, 85, 114, 129-130, 133, 135, 139
descodificación, en fonemas y sílabas: 128
descripción, niveles de: 50, 53
descubrimiento, interpretación y: 184-185; método de: 24, 181
designación, significación y: 189-192
desinencias: 58
diacríticos: 125-126
diagramas: 122
dialectos: 25, 187
Diccionario de Autoridades: 102
Diccionario del español de México: 149
diccionarios: 18-19, 133, 143, 147; amerindios: 11; entradas homonímicas en: 204; etimológicos: 238; de sinónimos: 202
disfemismos: 217, 219, 225, 228
dispersión de un vocablo: 167
disponibilidad léxica: 173; medidas de la: 175
doble articulación, teoría de la: 58-59

Edad Media: 80, 102, 127, 204, 232; Alta: 128, 130, 234; Baja: 18
egipcia, escritura: 115
enclítica, posición: 65
energeia: 39-40, 88
ergon: 40, 88
escritura: 19; alfabética: 65, 112; configuración morfológica: 121-123; continua: 129-130; demótica: 26; fonográfica: 115-127; glotográfica: 113-114, 118; ideográfica: 65, 112-115; jeroglífica: 112, 115; orden lineal y dirección: 118-119; principales características de la: 112; pronunciación y: 127; reglas de complexión: 120-127; reglas de representación: 118; silábica: 65; sistema de: 113-115; trazos: 116-118
España, purismo: 226
español, función distintiva del acento en: 59-60
esquemas, de acción: 92; de conocimiento: 92, 99; de conocimiento senso-motor: 44, 86-88; formación de: 88; morfotácticos: 65-66, 81-82; oracionales: 59, 111;

signo y: 93; silábicos: 65; secuenciales: 65
estado de lengua, sincronía y: 185-187
estereotipo(s): 93-99, 103, 105, 111; características del: 99; formación de: 102
estratos del significado: 92, 96, 101, 105
estructura, del léxico: 181, 204; sémica de los vocablos: 196-197; silábica: 30-31
estructuralismo: 192, 199
étimo: 233-234
etimología, culta: 242; etimología de: 232; *ex causa*: 232; *ex origine*: 233; como historia de palabras: 231, 233; popular: 47, 204, 241-243
eufemismos: 214, 216-219, 221, 225, 228; definición: 215
experiencia, histórica: 99; sensorial: 86-87, 106; *véase también* sentidos
expresión, contenido y: 53-57; forma de la: 23, 29; del signo: 25, 34; sustancia de la: 23, 25-26, 56-57
extranjerismos: 222-223

Feijoo, Fray Benito Jerónimo: 227
fenicios, escritura: 116
figuras de pensamiento, *véase* tropos
filosofía, griega: 37; idealista: 54-55
fonema(s): 21-23, 59, 111, 147; cadenas de: 29, 58, 112; codificación de: 121; combinación entre: 58; fricativo: 33; con función demarcativa: 31-34, 63, 79; identificación de los: 23-28; negativos: 33; palabra como combinación de: 29; relación sílaba: 26-27; unidad: 17
fonética(s), experimental: 54; inspección: 21; leyes: 234; restricción: 34
fonografía: 125
fonología diacrónica: 231
fonológica(o), análisis: 26-34, 184; cambio: 234; complexión: 53, 70; formación: 31; oposición: 32
fonosimbolismo: 239-241
forma(s), de la expresión: 25, 29; del contenido: 57; relación sonido y: 55; supletivas: 71-72; sustancia y: 53-57
Forner, Juan Pablo, *Exequias de la lengua castellana*: 226
Förstermann, Ernst: 242
francés, acento de intensidad: 33-34; análisis léxico: 205; escritura: 122-124; fundamental: 175; léxico: 173; onomástica: 244; onomatopeya: 241; purismo: 226
frecuencia de un vocablo: 167-168; absoluta y relativa: 169-170; porcentajes de: 169
función demarcativa: 70; afonológica: 32; fenómenos prosódicos con: 33-34; fonológica: 32; tres criterios de: 31-32
fuzzy logics (lógica difusa): 177

galicismos: 227-228
García Hoz, Víctor: 171-172
géneros discursivos o textuales, ponderación de: 169-171
germanías: 220-221, 225, 242
gerundio: 218
gestalt: 88, 90-91, 93, 96, 99; *véase también* psicología de la forma
gestalten: 90, 117
Gougenheim, Georges: 172-173, 175
grabación: 160
grafía: 117
gráficas, construcciones: 24
gramática(s): 18, 133; amerindia: 11; definición de: 39; generativo-transformacional: 198-199; griega: 37-38; histórica: 231, 234; latina: 38, 128-129; medieval: 131; náhuatl: 136
gramemas: 64, 67, 71; derivativo: 81; de género: 81; de número: 81; preposicional: 81, 94
Greenberg, Joshua: 68
griego, corpus del, antiguo: 157; escritura: 114
groserías: 221, 225

habla, distinción entre lengua y: 22
Hallig, Rudolf, *Begriffssystem als Grundlage für die Lexikographie*: 207
hapax legomena: 152-153, 162
hebreo: 116; moderno: 148; neologismos: 224
Heger, Klaus: 201
Hermógenes: 232
hiperonimia: 201-203
hipocorísticos: 46, 240
Hjelmslev, Louis: 23, 53-54, 56, 85, 109
homófonos: 125-126
homógrafos: 120
homonimia: 106, 203-205; definición: 107
Humboldt, Alejandro de: 40
Humboldt, Guillermo de: 40
húngaro, análisis léxico: 205; fonosimbolismo: 240

ideogramas: 148
ideologías lingüísticas y sociales: 228-229
índice de disponibilidad léxica: 175
información sensorial: 86; *véase también* sentidos
informatividad, y redundancia: 123-125
inglés, fonosimbolismo: 240; onomatopeya: 241;verbos: 94-95
inteligencia, formación de la: 88, 111
interjección, exclamativa: 125
interpretación, descubrimiento e: 184-185
invariantes: 25, 55; búsqueda de las: 22-23
Isidoro de Sevilla, San: 233, 235; *Etymologiae*: 232, 241
italiano, fonosimbolismo: 240; onomástica: 244

Jakobson, Roman, *Studies on Child Language and Aphasia*: 44
japonés, patrón silábico del: 31
Jenofonte: 152
jergas: 220-221, 225
jeroglíficos, egipcios: 26; mayas: 27
Jespersen, Otto: 240
Jóvenes Gramáticos, escuela: 234
Juilland, Alphonse: 169; *A Frequency Dictionary of Spanish Words*: 168

Landa, Diego de: 26
latín, acento de intensidad: 34; comparado con el castellano: 38; corpus: 157; declinaciones: 139; escritura: 114; fonosimbolismo: 240; neologismos: 224; paradigmas cerrados: 128; verbos: 139; vulgar: 234
lectura: 128; psicología de la: 130
lengua(s), amerindias: 11, 39, 79, 137, 148, 157-158, 245; clásicas: 38; comparación entre: 205-206; distinción entre habla y: 22; doble articulación de la: 58; estado de: 185-187; histórica: 55, 96, 98, 100, 103, 111, 147, 149, 151-153, 158-159, 161, 186; métodos de descripción de: 24;

modernas: 38; muertas: 157; como objeto de estudio: 39; romances: 235; vulgares: 38; *véase también* las distintas lenguas
letras: 117
lexemas: 65, 67-68, 72, 94; ligado: 81; paradigma de los: 64, 71
léxico, análisis semántico del: 181, 185, 188, 191; campo: 197; carácter ilimitado del: 147; efectos ideológicos sobre el: 225-229; enfoque cualitativo del: 176; estructura del: 181; estudios cuantitativo del: 165; fenómenos del: 107, 145; de una lengua histórica: 151-153; memoria y: 143, 174, 178, 181; relación con la memoria humana: 12; socialidad del: 174; de solidaridad: 219-221, 225; tres dimensiones del: 144-145; *véase también* vocabulario
lexicografía: 144
lexicología, como quinto nivel de análisis: 82
lingüística, como ciencia empírica: 9; comparativa: 233; condición de posibilidad de la: 22; descriptiva: 37, 48-49, 135; espacial: 235; estructural: 185, 233; histórica: 233; innatista: 54-55; pragmática: 40, 192; del producto: 40
Lope Blanch, Juan M.: 238
López Chávez, Juan: 175; *¿Qué te viene a la memoria?*: 174

Malkiel, Yakov: 235, 237-238
marcadores semánticos: 198-199
Martinet, André: 18-19, 58-59, 64; *La lingüística. Guía alfabética*: 17, 37
masculino genérico: 136
materia sonora: 25, 54-55; continua: 56
maya: 27; escritura: 26, 115; paradigmas de clasificadores semánticos: 210
Melchuk, Igor: 196
memoria: 151; colectiva: 143, 178, 183; individual: 143, 181; de la lengua histórica: 158; léxico y: 143, 174, 178, 181; social: 181, 197
metáfora, formación de la: 111; y la polisemia: 108-109
métodos, de la conmutación: 24-27, 181-182, 184, 199; de descubrimiento: 24, 181; interpretativo: 185; de pares mínimos: 25, 60-61, 182, 184
metonimia: 108
México, diversidad de lenguas en: 11
micénica, escritura: 114
minoica, lengua: 114
mixteco, paradigmas de clasificadores semánticos: 210
monema: 18
morfema(s): 9, 59-61, 111, 147; absolutivos: 136; anomalías entre: 71; cadenas de: 68; cohesión entre: 67-71, 77-80, 82, 85, 128, 130, 137; definición de: 62; derivativos: 64, 67, 136; de diminutivo: 33, 67; especificadores: 139; de flexión: 64, 67-68, 136; de género: 134; de inventario abierto: 64; de inventario cerrado: 64, 125, 131; léxico: 125, 131, 138; libres y ligados: 63-67, 76, 138; ligados con cohesión débil: 131; ligados de inventarios cerrados: 77-80, 85; de modo: 137; de número: 134, 137; orden de: 80-81; paradigmas de: 64-65, 111, 130-131; de persona: 137; posesivo: 50, 134; secuencia de: 79; segmentación de: 66-68, 70, 77, 82, 135, 137; temporales: 71, 137; unidad: 17
morfografía: 125
morfología, diacrónica: 231; y sintaxis: 57-58, 82
morfológico, cambio: 234-235
morfotáctica: 65-77, 80; definición: 65; esquemas: 65-66
Mosterín, Jesús: 119
muestra de electores: 154

náhuatl, análisis léxico: 205; corpus: 157; etimologías: 238; fonosimbolismo: 240; gramática: 136; morfemas: 77; paradigmas de los sustantivos: 64; préstamos al español: 238; reivindicación del: 76; transcripción: 158-159; unidades de denominación: 47, 50, 72; verbos: 137
Nebrija, Elio Antonio de, *Gramática de la lengua castellana*: 38
neologismos: 223-224
niveles de análisis: 21, 26, 27, 50, 80, 82
noemas: 208-210
nombres, motivación de los: 215-216, 232
nomenclaturas: 190
núcleo morfemático: 68-70, 78, 131, 137-138; con formas supletivas: 71-72; secuencias de: 72-75

objetivación: 40; como condición para la lingüística: 39; del idioma: 41; de los signos: 41
ocurrencia de vocablos: 155
onomasiología: 205-210
onomástica: 243-246
onomatopeya, definición: 240-241

oración: 38, 58, 67, 147; esquemas de la: 59, 111; partes de la: 39, 128
oralidad: 112, 131
organización por niveles: 19 y ss.
órganos fonatorios: 53-54
ortografía(s): 18, 126; definición: 127; reglas de: 218

pachuco: 220
palabra(s), agudas: 120; como combinación de fonemas: 29; composiciones de: 74; concepto gráfico de: 19; determinación de la unidad: 35; distinción entre vocablo y: 135; esdrújula: 33, 120; explicación lingüística de: 18; fonológica: 34-35, 37; frontera de la: 31, 34; gramaticales: 78-80; graves: 120; historia de: 231, 233; introducción de espacio entre: 130; con más de un significado: 106, 108-109; morfología de la: 53; ocurrencia de: 155-157, 167; percepción de la, escrita: 130; primera condición para determinar la: 28; reflexión en la determinación de la: 134-135; relacionales: 80; representación de la: 133; segmentación de: 130; como símbolos sociales: 213; sustancia de contenido de la: 85; teoría de la: 12; tipo: 155-157; tipo I: 63, 72, 78, 135; tipo II: 70, 72, 78, 135; tipo III: 75, 78, 135; tipo IV: 80, 131, 137; tipos de unidad: 82; vocablo y: 138-139
Palenque, inscripciones en: 27
paleografía: 117
paradigmas, cerrados: 128, 138, 210; de clasificadores semánticos: 210; de

conjugación de verbos: 138; de declinación: 135; derivacional: 135, 138; flexional: 135, 138, 140; con formas supletivas: 72; de morfemas: 64-65, 78, 111, 130-131
pares mínimos, método de: 25, 60-61, 182
Parkes, Malcolm B.: 128-129
perífrasis: 208; de futuro: 203
permutaciones: 29
pertinencia del rasgo: 25, 193
physei, doctrina de la: 216
Piaget, Jean: 44-45, 55; sobre la formación de la inteligencia: 88, 111
piedra de Roseta: 26
plano de la expresión, *véase* significante
Platón: 216; *Cratilo*: 38, 215, 232, 239, 241, 243
polisemia: 107, 111, 205, 207; definición: 106; metáfora y: 108-109
Pottier, Bernard: 196
predicado: 128
prefijos: 65; preposicionales: 94
preposiciones: 79, 94, 137, 218
préstamos de otras lenguas: 222, 225, 227, 236-239; *véase también* extranjerismos
principio fonológico, definición: 118
proclítica, posición: 65
producto, actividad y: 39-41; lingüística del: 40
pronombres: 137
pronunciación, y escritura: 127
prototipo, definición: 91-92
prueba de intercalabilidad: 67
psicología, experimental: 92, 111; de la forma (de la Gestalt): 88; de la lectura: 130
purismo: 226-227, 229
Putnam, Hilary: 96
raíces: 58; verbales: 68
rasgos significativos, *véase* semas
Real Academia de la Lengua: 102, 162
redundancia, informatividad y: 123-125
referencia, diferencia entre significado y: 231
representación, reglas de: 115, 118
representatividad: 157; cálculos de: 159; problema de la: 154-155
retórica, antigua: 108-109
Rieger, Burghard: 176, 183
Rosch, Eleanor: 90-91
Rosenblat, Ángel: 218
ruso: 117; análisis léxico: 205

Sagan, Carl: 20
Sampson, Geoffrey: 125
Sapir: 18-19
Saussure, Ferdinand de: 22-24, 39-40, 53, 104, 205
Schleicher, August: 236
secuencias de núcleos morfemáticos: 72
segmentación, hipótesis de: 76; de morfemas: 66-67, 70, 82, 135, 137; de palabras: 130; riesgos de la: 75-77; del texto: 129
semántica: 57; estructural: 192, 196, 198
semántica léxica: 181, 216; métodos de investigación de la: 12
sema(s)s: 182, 195, 197; características de un: 192-193; definición: 193; formulación del: 196-197
semasiología: 206-210; definición: 205
semema: 195-197
semítica, lengua: 116
sentidos: 85, 102; configuración perceptual: 86-92; información de los: 86
significación, designación y: 189-192
significado: 23, 25, 53; construcción del: 85-106; cuarto estrato de la formación del: 105; la cultura y la formación del: 99-103, 111; definición de: 37; descomposición interpretativa del: 196; diferencia entre referencia y: 231; especializado: 103-105; como inseparable del significante: 24; primer estrato de la formación del: 92; reciprocidad del: 60; segundo estrato de formación del: 96; en el segundo nivel de análisis: 24; tercer estrato de formación del: 101-102; valor de apoyo del: 24
significante: 23, 25, 37, 53; como inseparable del significado: 24; reciprocidad del: 60
signo(s), análisis estructural de los: 181; conmutación entre: 40; destrucción del: 24; esquema y: 93; forma de expresión del: 25, 34; funciones bühlerianas del: 201; identificación de los: 26; objetivación de los: 41; plano del contenido del: 24; significado de los: 23; teoría saussureana del: 23-24, 37, 48, 53, 205, 207, 231
sílaba(s), esquemas de las: 65; estructura de las: 31; margen de la: 30; núcleo de la: 30; patrón canónico de la: 30; patrones de las: 111; relación fonema: 26-27
símbolos sociales, efecto normativo de los: 216-225; palabras como: 213
sincategoremas: 80
sincronía, estado de lengua y: 185-187
sinécdoque: 108
sinonimia: 106, 202; apelativa: 201; definición: 107; notación de: 199; referencial: 200-201; sintomática: 201
sinónimos: 199
sintagma(s): 9, 18, 70; relativamente fijos: 75
sintaxis: 218; morfología y: 57-58, 82
sistema de escritura, ideográfico: 113; glotográfico: 113
Sócrates: 215-216, 232-233, 239-240, 243
solecismo: 224-225
sonidos, consonánticos: 22; identidad de los: 22; percepción de los: 54; relación forma y: 55; vocálicos: 22
Steinthal, Hugo: 201
sufijos: 65; de número: 124
sujeto: 128
sustancia, del contenido: 57; de la expresión: 23, 25-26, 56-57; forma y: 53-57
sustantivo(s): 38, 50, 67, 128; derivaciones del: 138; femenino: 137; flexiones del: 138; masculino: 134; paradigma de los: 64; patrones más frecuentes en: 81
sustitución paradigmática: 72; capacidad de: 70

tabú verbal: 214-216
teoría, de la doble articulación: 58-59; saussureana del signo: 23-24, 37, 48, 53, 205, 297, 231
término técnico, estrato del: 105
términos científicos y técnicos: 190
tertium comparationis: 206-207
texto(s): 161; segmentación del: 129
thesei, doctrina de la: 216
tipo de vocablos: 155
Tibón, Gutierre, *Diccionario etimológico comparado de nombres propios de persona*: 243
toponimia: 244-245

topónimos: 243
tradición(es), culta: 149, 188; popular: 149, 188, 214; verbales: 187-188, 214, 216
transcripción: 159; fina y gruesa: 157-158
trazos: 115-118
Trésor de la langue française: 156
trigrama: 122
trivium: 232
tropos: 108
tzeltal: 28, 30; acento de intensidad: 34; análisis léxico: 205; corpus de datos: 153; morfemas: 64; unidades de denominación: 61; verbos: 95, 137

Unamuno, Miguel de: 227
unidad dialectal: 187
unidades de cita: 49-50, 133-135, 139; características de las: 135-137; papel de la: 137-138; de primera articulación: 61
unidades de denominación: 48, 61-63, 66, 69-70, 72-75, 78-79, 85, 114, 139; en alemán: 47; en náhuatl: 47, 50; obtención de: 48-50; reconocimiento de: 129-130, 133, 135; segmentación de: 62

vaguedad: 106, 108; confusión entre ambigüedad y: 107
valor, de campo: 182; denominativo: 53, 70
verbos: 38, 50, 65, 67, 72, 93, 95, 139; conjugaciones: 128, 138; esquema de acción: 100; infinitivo: 134, 137; morfología de: 94
vocablo(s): 133, 155-157; constitución histórica del: 139-140; definición: 138; dispersión de un: 167-168, 170; distinción entre palabra y: 135; estructura sémica de los: 196; frecuencia de un: 167-168; nominales: 161; ocurrencia de: 168, 171; palabra y: 138-139; significado de los: 145; tipo de: 155
vocabulario, activo: 149-151; común: 171-172; disponible: 150, 164, 172-174, 181, 183; estudios etimológicos del: 145; fundamental: 147-149, 164, 171-173; pasivo: 150-151; *véase también* léxico
vocales, abiertas: 122-124; agudas: 240; anteriores: 28; cerradas: 122-124; media: 28; nasales: 32; oscuras: 240; posteriores: 28; representación de las: 116; temáticas: 67-68, 71
Voltaire, François-Marie Arouet: 233

Wartburg, Walter von, *Begriffssystem als Grundlage für die Lexikographie*: 207; *Französisch Etymologisches Wörterbuch*: 236
Wierzbicka, Anna: 196, *Lingua mentalis: The Semantics of Natural Language*: 207
Worter und Sachen, escuela: 206

Zamboni, Alberto: 236
Zamudio, Cecilia: 159
zapoteco, análisis léxico: 205
Zauner, Adolf, *Die romanische Namen der Körperteile*: 206
zoque, corpus de datos: 153

Curso de lexicología
se terminó de imprimir en enero de 2006
en los talleres de Editorial Color, S.A. de C.V.,
Naranjo 96 bis PB, Sta. María la Ribera, 06400 México, D.F.
Portada de Roberto Campos.
Tipografía y formación: Socorro Gutiérrez,
en Redacta, S.A. de C.V.
Cuidaron la edición Eugenia Huerta y el autor.